OPIC
OFFICE DE LA PROPRIÉTÉ INTELLECTUELLE DU CANADA

CIPO
CANADIAN INTELLECTUAL PROPERTY OFFICE

(12) (19) (CA) **Demande-Application**

(21) (A1) **2,271,718**
(86) 1997/11/13
(87) 1998/05/22

(72) HARRINGTON, Lea A., CA
(72) ROBINSON, Murray O., US
(71) AMGEN INC., US
(71) AMGEN CANADA INC., CA

(51) Int.Cl.[6] C12N 15/54, C12N 15/85, A61K 31/70, A01K 67/027, C12Q 1/68, C12Q 1/48, A61K 38/45, C07K 16/40, C12N 1/21, C12N 1/19, C12N 9/12, C12N 15/11
(30) 1996/11/15 (08/751,189) US
(30) 1997/06/11 (08/873,039) US
(30) 1997/10/16 (08/951,733) US

(54) **GENES CODANT DES PROTEINES DE TELOMERASE**
(54) **GENES ENCODING TELOMERASE PROTEINS**

(57) L'invention concerne des molécules d'acide nucléique, qui codent des polypeptides du complexe télomérase. L'invention se rapporte également à des procédés de préparation desdites molécules d'acide nucléique et desdits polypeptides et à des procédés d'utilisation desdites molécules.

(57) Disclosed are nucleic acid molecules encoding polypeptides of the telomerase complex. Also disclosed are methods of preparing the nucleic acid molecules and polypeptides, and methods of using these molecules.

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/54, 9/12, C12Q 1/68, 1/48, C12N 15/11, 15/85, A01K 67/027, C07K 16/40, A61K 38/45, 31/70, C12N 1/21, 1/19

A1

(11) International Publication Number: WO 98/21347

(43) International Publication Date: 22 May 1998 (22.05.98

(21) International Application Number: PCT/US97/21248

(22) International Filing Date: 13 November 1997 (13.11.97)

(30) Priority Data:

08/871,189	15 November 1996 (15.11.96)	US
08/873,039	11 June 1997 (11.06.97)	US
08/951,733	16 October 1997 (16.10.97)	US

(71) Applicants: AMGEN INC. [US/US]; Amgen Center, 1840 De Havilland Drive, Thousand Oaks, CA 91320-1789 (US). AMGEN CANADA INC. [CA/CA]; Suite 303, 6733 Mississauga Road, Mississauga, Ontario L5N 6J5 (CA).

(72) Inventors: HARRINGTON, Lea, A.; 55 Pears Avenue, Toronto, Ontario M5R 1S9 (CA). ROBINSON, Murray, O.; 22623 Pacific Coast Highway, Malibu, CA 90265 (US).

(74) Agents: ODRE, Steven, M. et al.; Amgen, Inc., Amgen Center, 1840 De Havilland Drive, Thousand Oaks, CA 91320-1789 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BF BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ BY, KG, KZ, MD, RU, TJ, TM) European patent (AT, BE CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN ML, MR, NE, SN, TD, TG).

Published
With international search report.
With amended claims and statement.

Date of publication of the amended claims and statement:
30 July 1998 (30.07.9

(54) Title: GENES ENCODING TELOMERASE PROTEINS

(57) Abstract

Disclosed are nucleic acid molecules encoding polypeptides of the telomerase complex. Also disclosed are methods of preparing th nucleic acid molecules and polypeptides, and methods of using these molecules.

AMENDED CLAIMS

[received by the International Bureau on 19 June 1998 (19.06.98); new claims 33-56 added; remaining claims unchanged (7 pages)]

1. A TP2 nucleic acid molecule encoding a polypeptide selected from the group consisting of:

(a) the nucleic acid molecule of SEQ ID NO:13;

(b) the nucleic acid molecule that is nucleotides 1920-2820 of SEQ ID NO:13;

(c) the nucleic acid molecule of SEQ ID NO:19

(d) a nucleic acid molecule encoding the polypeptide of SEQ ID NO:14, or a biologically active fragment thereof;

(e) a nucleic acid molecule encoding the polypeptide of SEQ ID NO:20, or a biologically active fragment thereof;

(f) a nucleic acid molecule that encodes a polypeptide that is at least 90 percent identical to the polypeptide of SEQ ID NO:14;

(g) a nucleic acid molecule that encodes a polypeptide that is at least 90 percent identical to the polypeptide of SEQ ID NO:20;

(h) a nucleic acid molecule that hybridizes under stringent conditions to any of (a)-(g) above; and

(i) a nucleic acid molecule that is the complement of any of (a)-(g) above.

2. The nucleic acid molecule that is SEQ ID NO:13 or SEQ ID NO:19.

3. The nucleic acid molecule that is nucleotides 1920-2820 of SEQ ID NO:13.

4. A nucleic acid molecule encoding the polypeptide of SEQ ID NO:14 of SEQ ID NO:20.

5. A nucleic acid molecule selected from the group consisting of: nucleotides 1-1689 of SEQ ID NO:13, nucleotides 1-1920 of SEQ ID NO:13, nucleotides 1920-2820 of SEQ ID NO:13, nucleotides 2089-2820 of SEQ ID NO:13, and nucleotides 2089-2859 of SEQ ID NO:13.

6. A nucleic acid molecule encoding amino acids 640-940 of the polypeptide of SEQ ID NO:14.

7. A vector comprising the nucleic acid molecule of claim 1.

8. A vector comprising the nucleic acid molecule of claim 2.

9. A vector comprising the nucleic acid molecule of claim 3.

10. A vector comprising the nucleic acid molecule of claim 4.

11. A vector comprising the nucleic acid molecule of claim 5.

12. A vector comprising the nucleic acid molecule of claim 6.

13. A host cell comprising the vector of claim 7.

14. A host cell comprising the vector of claim 8.

15. A host cell comprising the vector of claim 9.

16. A host cell comprising the vector of claim 10.

17. A host cell comprising the vector of claim 11.

18. A host cell comprising the vector of claim 12.

19. A process for producing a TP2 polypeptide comprising the steps of:

(a) expressing a polypeptide encoded by the nucleic acid of claim 1 in a suitable host; and

(b) isolating the polypeptide.

20. The process of claim 19 wherein the polypeptide is SEQ ID NO:14 or SEQ ID NO:20.

21. The process of claim 19 wherein the polypeptide is amino acids 640-940 of SEQ ID NO:14.

22. A TP2 polypeptide selected from the group consisting of:

(a) the polypeptide of SEQ ID NO:14;

(b) the polypeptide that is amino acids 640-940 of SEQ ID NO:14;

(c) the polypeptide of SEQ ID NO:20; and

(d) a polypeptide that is at least 90 percent identical to any of the polypeptides of (a)-(c).

23. A TP2 polypeptide that is the polypeptide of SEQ ID NO:14, SEQ ID NO:20, or a biologically active fragment thereof.

24. A TP2 polypeptide selected from the group consisting of: amino acids 1-563 of SEQ ID NO:14; amino acids 1-640 of SEQ ID NO:14; amino acids 640-940 of SEQ ID NO:14; amino acids 696-940 of SEQ ID NO:14; and amino acids 696-953 of SEQ ID NO:14.

25. The TP2 polypeptide of claim 22 that does not possess an amino terminal methionine.

26. A method of increasing proliferation of a cell, comprising expressing a nucleic acid encoding TP2 or a biologically active fragment thereof, in the cell.

27. A method of increasing telomerase activity in a cell, comprising expressing a TP2 gene, or a biologically active fragment thereof, in the cell.

28. A method of decreasing telomerase in a cell, comprising expressing a TP2 mutant in a cell, wherein the mutant does not have TP2 biological activity.

29. A nucleic acid molecule encoding a mutant TP2 polypeptide, wherein the codon for aspartic acid at amino acid position 868 or 869 is changed to a codon for alanine.

30. A nucleic acid molecule encoding a mutant TP2 polypeptide, wherein the codons for aspartic acid at amino acid positions 868 and 869 are changed to codons for alanine.

31. A polypeptide encoded by the nucleic acid molecule of claim 29.

32. A polypeptide encoded by the nucleic acid molecule of claim 30.

33. A TRIP1 nucleic acid molecule encoding a polypeptide selected from the group consisting of:

(a) the nucleic acid molecule of SEQ ID NO:1;

(b) the nucleic acid molecule of SEQ ID NO:2;

(c) a nucleic acid molecule encoding the polypeptide of SEQ ID NO:3, SEQ ID NO:4, or a biologically active fragment thereof;

(d) a nucleic acid molecule that encodes a polypeptide that is at least 70 percent identical to the polypeptide of SEQ ID NO:3 or SEQ ID NO:4;

(e) a nucleic acid molecule that hybridizes under stringent conditions to any of (a)-(d) above; and

(f) a nucleic acid molecule that is the complement of any of (a)-(e) above.

34. The nucleic acid molecule that is SEQ ID NO:1.

35. The nucleic acid molecule that is SEQ ID NO:2.

36. A nucleic acid molecule encoding the polypeptide of SEQ ID NO:3.

37. A nucleic acid molecule encoding the polypeptide of SEQ ID NO:4.

38. A nucleic acid molecule encoding amino acids 1-871 of the polypeptide of SEQ ID NO:3.

39. A vector comprising the nucleic acid molecule of claim 33.

40. A vector comprising the nucleic acid molecule of claim 34.

41. A vector comprising the nucleic acid molecule of claim 35.

42. A vector comprising the nucleic acid molecule of claim 36.

43. A vector comprising the nucleic acid molecule of claim 37.

44. A vector comprising the nucleic acid molecule of claim 38.

45. A host cell comprising the vector of claim 39.

46. A host cell comprising the vector of claim 40.

47. A host cell comprising the vector of claim 41.

48. A host cell comprising the vector of claim 42.

49. A host cell comprising the vector of claim 43.

50. A host cell comprising the vector of claim 44.

51. A process for producing a TRIP1 polypeptide comprising the steps of:

(a) expressing a polypeptide encoded by the nucleic acid of claim 1 in a suitable host; and

(b) isolating the polypeptide.

52. The process of claim 51 wherein the polypeptide is SEQ ID NO:3.

53. The process of claim 51 wherein the polypeptide amino acids 1-871 of SEQ ID NO:3.

54. A TRIP1 polypeptide selected from the group consisting of:

(a) the polypeptide of SEQ ID NO:3;

(b) the polypeptide that is amino acids 1-871 of SEQ ID NO:3; and

(c) a polypeptide that is at least 70 percent identical to the polypeptide of (a) or (b).

55. A TRIP1 polypeptide that is the polypeptide of SEQ ID NO:3 or a biologically active fragment thereof.

56. The TRIP1 polypeptide of claim 52 that does not possess an amino terminal methionine.

STATEMENT UNDER ARTICLE 19

The claims of International Application WO 98/21248, published 22 May 1998, have been amended. Original claims 1 through 32 have not been amended, however, new claims 33 through 56 have been added. Claims 33 through 56 are directed to an aspect of the invention not originally claimed by Applicants. Specifically, claims 33 through 56 encompass telomerase protein 1 and DNA encoding therefor. Such claims are fully supported by the written description and the drawings.

FIG. 1A

ATGGAAAAACTCCATGGGCATGTGTCTGCCCATCCAGACATCCTCTCCT

TGGAGAACCGGTGCCTGGCTATGCTCCCTGACTTACAGCCCTTGGAGAA

ACTACATCAGCATGTATCTACCCACTCAGATATCCTCTCCTTGAAGAAC

CAGTGCCTAGCCACGCTTCCTGACCTGAAGACCATGGAAAAACCACATG

GATATGTGTCTGCCCACCCAGACATCCTCTCCTTGGAGAACCAGTGCCT

GGCCACACTTTCTGACCTGAAGACCATGGAGAAACCACATGGACATGTT

TCTGCCCACCCAGACATCCTCTCCTTGGAGAACCGGTGCCTGGCCACCC

TCCCTAGTCTAAAGAGCACTGTGTCTGCCAGCCCCTTGTTCCAGAGTCT

ACAGATATCTCACATGACGCAAGCTGATTTGTACCGTGTGAACAACAGC

AATTGCCTGCTCTCTGAGCCTCCAAGTTGGAGGGCTCAGCATTTCTCTA

AGGGACTAGACCTTTCAACCTGCCCTATAGCCCTGAAATCCATCTCTGC

CACAGAGACAGCTCAGGAAGCAACTTTGGGTCGTTGGTTTGATTCAGAA

GAGAAGAAAGGGGCAGAGACCCAAATGCCTTCTTATAGTCTGAGCTTGG

GAGAGGAGGAGGAGGTGGAGGATCTGGCCGTGAAGCTCACCTCTGGAGA

CTCTGAATCTCATCCAGAGCCTACTGACCATGTCCTTCAGGAAAAGAAG

ATGGCTCTACTGAGCTTGCTGTGCTCTACTCTGGTCTCAGAAGTAAACA

TGAACAATACATCTGACCCCACCCTGGCTGCCATTTTTGAAATCTGTCG

TGAACTTGCCCTCCTGGAGCCTGAGTTTATCCTCAAGGCATCTTTGTAT

GCCAGGCAGCAGCTGAACGTCCGGAATGTGGCCAATAACATCTTGGCCA

FIG. 1B

TTGCTGCTTTCTTGCCGGCGTGTCGCCCCCACCTGCGACGATATTTCTG

TGCCATTGTCCAGCTGCCTTCTGACTGGATCCAGGTGGCTGAGCTTTAC

CAGAGCCTGGCTGAGGGAGATAAGAATAAGCTGGTGCCCCTGCCCGCCT

GTCTCCGTACTGCCATGACGGACAAATTTGCCCAGTTTGACGAGTACCA

GCTGGCTAAGTACAACCCTCGGAAGCACCGGGCCAAGAGACACCCCCGC

CGGCCACCCCCGCTCTCCAGGGATGGAGCCTCCATTTTCTCACAGATGTT

TTCCAAGGTACATAGGGTTTCTCAGAGAAGAGCAGAGAAAGTTTGAGAA

GGCCGGTGATACAGTGTCAGAGAAAAAGAATCCTCCAAGGTTCACCCTG

AAGAAGCTGGTTCAGCGACTGCACATCCACAAGCCTGCCCAGCACGTTC

AAGCCCTGCTGGGTTACAGATACCCCTCCAACCTACAGCTCTTTTCTCG

AAGTCGCCTTCCTGGGCCTTGGGATTCTAGCAGAGCTGGGAAGAGGATG

AAGCTGTCTAGGCCAGAGACCTGGGAGCGGGAGCTGAGCCTACGGGGGA

ACAAAGCGTCGGTCTGGGAGGAACTCATTGAAAATGGGAAGCTTCCCTT

CATGGCCATGCTTCGGAACCTGTGCAACCTGCTGCGGGTTGGAATCAGT

TCCCGCCACCATGAGCTCATTCTCCAGAGACTCCAGCATGGGAAGTCGG

TGATCCACAGTCGGCAGTTTCCATTCAGATTTCTTAACGCCCATGATGC

CATTGATGCCCTCGAGGCTCAACTCAGAAATCAAGCATTGCCCTTTCCT

TCGAATATAACACTGATGAGGCGGATACTAACTAGAAATGAAAAGAACC

GTCCCAGGCGGAGGTTTCTTTGCCACCTAAGCCGTCAGCAGCTTCGTAT

FIG.1C

GGCAATGAGGATACCTGTGTTGTATGAGCAGCTCAAGAGGGAGAAGCTG

AGAGTACACAAGGCCAGACAGTGGAAATATGATGGTGAGATGCTGAACA

GGTACCGACAGGCCCTAGAGACAGCTGTGAACCTCTCTGTGAAGCACAG

CCTGCCCCTGCTGCCAGGCCGCACTGTCTTGGTCTATCTGACAGATGCT

AATGCAGACAGGCTCTGTCCAAAGAGCAACCCACAAGGGCCCCCGCTGA

ACTATGCACTGCTGTTGATTGGGATGATGATCACGAGGGCGGAGCAGGT

GGACGTCGTGCTGTGTGGAGGTGACACTCTGAAGACTGCAGTGCTTAAG

GCAGAAGAAGGCATCCTGAAGACTGCCATCAAGCTCCAGGCTCAAGTCC

AGGAGTTTGATGAAAATGATGGATGGTCCCTGAATACTTTTGGGAAATA

CCTGCTGTCTCTGGCTGGCCAAAGGGTTCCTGTGGACAGGGTCATCCTC

CTTGGCCAAAGCATGGATGATGGAATGATAAATGTGGCCAAACAGCTTT

ACTGGCAGCGTGTGAATTCCAAGTGCCTCTTTGTTGGTATCCTCCTAAG

AAGGGTACAATACCTGTCAACAGATTTGAATCCCAATGATGTGACACTC

TCAGGCTGTACTGATGCGATACTGAAGTTCATTGCAGAGCATGGGGCCT

CCCATCTTCTGGAACATGTGGGCCAAATGGACAAAATATTCAAGATTCC

ACCACCCCCAGGAAAGACAGGGGTCCAGTCTCTCCGGCCACTGGAAGAG

GACACTCCAAGCCCCTTGGCTCCTGTTTCCCAGCAAGGATGGCGCAGCA

TCCGGCTTTTCATTTCATCCACTTTCCGAGACATGCACGGGGAGCGGGA

CCTGCTGCTGAGGTCTGTGCTGCCAGCACTGCAGGCCCGAGCGGCCCCT

FIG.1D

CACCGTATCAGCCTTCACGGAATCGACCTCCGCTGGGGCGTCACTGAGG

AGGAGACCCGTAGGAACAGACAACTGGAAGTGTGCCTTGGGGAGGTGGA

GAACGCACAGCTGTTTGTGGGGATTCTGGGCTCCCGTTATGGATACATT

CCCCCCAGCTACAACCTTCCTGACCATCCACACTTCCACTGGGCCCAGC

AGTACCCTTCAGGGCGCTCTGTGACAGAGATGGAGGTGATGCAGTTCCT

GAACCGGAACCAACGTCTGCAGCCCTCTGCCCAAGCTCTCATCTACTTC

CGGGATTCCAGCTTCCTCAGCTCTGTGCCAGATGCCTGGAAATCTGACT

TTGTTTCTGAGTCTGAAGAGGCCGCATGTCGGATCTCAGAACTGAAGAG

CTACCTAAGCAGACAGAAAGGGATAACCTGCCGCAGATACCCCTGTGAG

TGGGGGGGTGTGGCAGCTGGCCGGCCCTATGTTGGCGGGCTGGAGGAGT

TTGGGCAGTTGGTTCTGCAGGATGTATGGAATATGATCCAGAAGCTCTA

CCTGCAGCCTGGGGCCCTGCTGGAGCAGCCAGTGTCCATCCCAGACGAT

GACTTGGTCCAGGCCACCTTCCAGCAGCTGCAGAAGCCACCGAGTCCTG

CCCGGCCACGCCTTCTTCAGGACACAGTGCAACAGCTGATGCTGCCCCA

CGGAAGGCTGAGCCTGGTGACGGGGCAGTCAGGACAGGGCAAGACAGCC

TTCCTGGCATCTCTTGTGTCAGCCCTGCAGGCTCCTGATGGGGCCAAGG

TGGCACCATTAGTCTTCTTCCACTTTTCTGGGGCTCGTCCTGACCAGGG

TCTTGCCCTCACTCTGCTCAGACGCCTCTGTACCTATCTGCGTGGCCAA

CTAAAAGAGCCAGGTGCCCTCCCCAGCACCTACCGAAGCCTGGTGTGGG

FIG. 1E

AGCTGCAGCAGAGGCTGCTGCCCAAGTCTGCTGAGTCCCTGCATCCTGG

CCAGACCCAGGTCCTGATCATCGATGGGGCTGATAGGTTAGTGGACCAG

AATGGGCAGCTGATTTCAGACTGGATCCCAAAGAAGCTTCCCCGGTGTG

TACACCTGGTGCTGAGTGTGTCTAGTGATGCAGGCCTAGGGGAGACCCT

TGAGCAGAGCCAGGGTGCCCACGTGCTGGCCTTGGGGCCTCTGGAGGCC

TCTGCTCGGGCCCGGCTGGTGAGAGAGGAGCTGGCCCTGTACGGGAAGC

GGCTGGAGGAGTCACCATTTAACAACCAGATGCGACTGCTGCTGGTGAA

GCGGGAATCAGGCCGGCCGCTCTACCTGCGCTTGGTCACCGATCACCTG

AGGCTCTTCACGCTGTATGAGCAGGTGTCTGAGAGACTCCGGACCCTGC

CTGCCACTGTCCCCCTGCTGCTGCAGCACATCCTGAGCACACTGGAGAA

GGAGCACGGGCCTGATGTCCTTCCCCAGGCCTTGACTGCCCTAGAAGTC

ACACGGAGTGGTTTGACTGTGGACCAGCTGCACGGAGTGCTGAGTGTGT

GGCGGACACTACCGAAGGGGACTAAGAGCTGGGAAGAAGCAGTGGCTGC

TGGTAACAGTGGAGACCCCTACCCCATGGGCCCGTTTGCCTGCCTCGTC

CAGAGTCTGCGCAGTTTGCTAGGGGAGGGCCCTCTGGAGCGCCCTGGTG

CCCGGCTGTGCCTCCCTGATGGGCCCCTGAGAACAGCAGCTAAACGTTG

CTATGGGAAGAGGCCAGGGCTAGAGGACACGGCACACATCCTCATTGCA

GCTCAGCTCTGGAAGACATGTGACGCTGATGCCTCAGGCACCTTCCGAA

GTTGCCCTCCTGAGGCTCTGGGAGACCTGCCTTACCACCTGCTCCAGAG

FIG. 1F

CGGGAACCGTGGACTTCTTTCGAAGTTCCTTACCAACCTCCATGTGGTG

GCTGCACACTTGGAATTGGGTCTGGTCTCTCGGCTCTTGGAGGCCCATG

CCCTCTATGCTTCTTCAGTCCCCAAAGAGGAACAAAAGCTCCCCGAGGC

TGACGTTGCAGTGTTTCGCACCTTCCTGAGGCAGCAGGCTTCAATCCTC

AGCCAGTACCCCCGGCTCCTGCCCCAGCAGGCAGCCAACCAGCCCCTGG

ACTCACCTCTTTGCCACCAAGCCTCGCTGCTCTCCCGGAGATGGCACCT

CCAACACACACTACGATGGCTTAATAAACCCCGGACCATGAAAAATCAG

CAAAGCTCCAGCCTGTCTCTGGCAGTTTCCTCATCCCCTACTGCTGTGG

CCTTCTCCACCAATGGGCAAAGAGCAGCTGTGGGCACTGCCAATGGGAC

AGTTTACCTGTTGGACCTGAGAACTTGGCAGGAGGAGAAGTCTGTGGTG

AGTGGCTGTGATGGAATCTCTGCTTGTTTGTTCCTCTCCGATGATACAC

TCTTTCTTACTGCCTTCGACGGGCTCCTGGAGCTCTGGGACCTGCAGCA

TGGTTGTCGGGTGCTGCAGACTAAGGCTCACCAGTACCAAATCACTGGC

TGCTGCCTGAGCCCAGACTGCCGGCTGCTAGCCACCGTGTGCTTGGGAG

GATGCCTAAAGCTGTGGGACACAGTCCGTGGGCAGCTGGCCTTCCAGCA

CACCTACCCCAAGTCCCTGAACTGTGTTGCCTTCCACCCAGAGGGGCAG

GTAATAGCCACAGGCAGCTGGGCTGGCAGCATCAGCTTCTTCCAGGTGG

ATGGGCTCAAAGTCACCAAGGACCTGGGGGCACCCGGAGCCTCTATCCG

TACCTTGGCCTTCAATGTGCCTGGGGGGGTTGTGGCTGTGGGCCGGCTG

FIG.1G

GACAGTATGGTGGAGCTGTGGGCCTGGCGAGAAGGGGCACGGCTGGCTG

CCTTCCCTGCCCACCATGGCTTTGTTGCTGCTGCGCTTTTCCTGCATGC

GGGTTGCCAGTTACTGACGGCTGGAGAGGATGGCAAGGTTCAGGTGTGG

TCAGGGTCTCTGGGTCGGCCCCGTGGGCACCTGGGTTCCCTTTCTCTCT

CTCCTGCCCTCTCTGTGGCACTCAGCCCAGATGGTGATCGGGTGGCTGT

TGGATATCGAGCGGATGGCATTAGGATCTACAAAATCTCTTCAGGTTCC

CAGGGGGCTCAGGGTCAGGCACTGGATGTGGCAGTGTCCGCCCTGGCCT

GGCTAAGCCCCAAGGTATTGGTGAGTGGTGCAGAAGATGGGTCCTTGCA

GGGCTGGGCACTCAAGGAATGCTCCCTTCAGTCCCTCTGGCTCCTGTCC

AGATTCCAGAAGCCTGTGCTAGGACTGGCCACTTCCCAGGAGCTCTTGG

CTTCTGCCTCAGAGGATTTCACAGTGCAGCTGTGGCCAAGGCAGCTGCT

GACGCGGCCACACAAGGCAGAAGACTTTCCCTGTGGCACTGAGCTGCGG

GGACATGAGGGCCCTGTGAGCTGCTGTAGTTTCAGCACTGATGGAGGCA

GCCTGGCCACCGGGGGCCGGGATCGGAGTCTCCTCTGCTGGGACGTGAG

GACACCCAAAACCCCTGTTTTGATCCACTCCTTCCCTGCCTGTCACCGT

GACTGGGTCACTGGCTGTGCCTGGACCAAAGATAACCTACTGATATCCT

GCTCCAGTGATGGCTCTGTGGGGCTCTGGGACCCAGAGTCAGGACAGCG

GCTTGGTCAGTTCCTGGGTCATCAGAGTGCTGTGAGCGCTGTGGCAGCT

GTGGAGGAGCACGTGGTGTCTGTGAGCCGGGATGGGACCTTGAAAGTGT

FIG.1H

GGGACCATCAAGGCGTGGAGCTGACCAGCATCCCTGCTCACTCAGGACC

CATTAGCCACTGTGCAGCTGCCATGGAGCCCCGTGCAGCTGGACAGCCT

GGGTCAGAGCTTCTGGTGGTAACCGTCGGGCTAGATGGGGCCACACGGT

TATGGCATCCACTCTTGGTGTGCCAAACCCACACCCTCCTGGGACACAG

CGGCCCAGTCCGTGCTGCTGCTGTTTCAGAAACCTCAGGCCTCATGCTG

ACCGCCTCTGAGGATGGTTCTGTACGGCTCTGGCAGGTTCCTAAGGAAG

CAGATGACACATGTATACCAAGGAGTTCTGCAGCCGTCACTGCTGTGGC

TTGGGCACCAGATGGTTCCATGGCAGTATCTGGAAATCAAGCTGGGGAA

CTAATCTTGTGGCAGGAAGCTAAGGCTGTGGCCACAGCACAGGCTCCAG

GCCACATTGGTGCTCTGATCTGGTCCTCGGCACACACCTTTTTTGTCCT

CAGTGCTGATGAGAAAATCAGCGAGTGGCAAGTGAAACTGCGGAAGGGT

TCGGCACCCGGAAATTTGAGTCTTCACCTGAACCGAATTCTACAGGAGG

ACTTAGGGGTGCTGACAAGTCTGGATTGGGCTCCTGATGGTCACTTTCT

CATCTTGGCCAAAGCAGATTTGAAGTTACTTTGCATGAAGCCAGGGGAT

GCTCCATCTGAAATCTGGAGCAGCTATACAGAAAATCCTATGATATTGT

CCACCCACAAGGAGTATGGCATATTTGTCCTGCAGCCCAAGGATCCTGG

AGTTCTTTCTTTCTTGAGGCAAAAGGAATCAGGAGAGTTTGAAGAGAGG

CTGAACTTTGATATAAACTTAGAGAATCCTAGTAGGACCCTAATATCGA

TAACTCAAGCCAAACCTGAATCTGAGTCCTCATTTTTGTGTGCCAGCTC

FIG. 1I

TGATGGGATCCTATGGAACCTGGCCAAATGCAGCCCAGAAGGAGAATGG

ACCACAGGTAACATGTGGCAGAAAAAAGCAAACACTCCAGAAACCCAAA

CTCCAGGGACAGACCCATCTACCTGCAGGGAATCTGATGCCAGCATGGA

TAGTGATGCCAGCATGGATAGTGAGCCAACACCACATCTAAAGACACGG

CAGCGTAGAAAGATTCACTCGGGCTCTGTCACAGCCCTCCATGTGCTAC

CTGAGTTGCTGGTGACAGCTTCGAAGGACAGAGATGTTAAGCTATGGGA

GAGACCCAGTATGCAGCTGCTGGGCCTGTTCCGATGCGAAGGGTCAGTG

AGCTGCCTGGAACCTTGGCTGGGCGCTAACTCCACCCTGCAGCTTGCCG

TGGGAGACGTGCAGGGCAATGTGTACTTTCTGAATTGGGAA

FIG.2A

ATGGAGAAGCTCTGTGGGCATGTGCCTGGCCATTCAGACATCCTCTCCT

TGAAGAACCGGTGCCTGACCATGCTCCCTGACCTCCAGCCCCTGGAGAA

AATACATGGACATAGATCTGTCCACTCAGACATCCTTTCCTTGGAGAAC

CAGTGTCTGACCATGCTCTCTGACCTCCAGCCCACGGAGAGAATAGATG

GGCATATATCTGTCCACCCAGACATCCTCTCCTTGGAGAATCGGTGCCT

GACCATGCTCCCTGACCTCCAGCCTCTGGAGAAGCTATGTGGACATATG

TCTAGTCATCCAGACGTCCTTTCTTTGGAAAACCAATGTCTAGCTACTC

TCCCCACTGTAAAGAGCACTGCATTGACCAGCCCCTTGCTCCAGGGTCT

TCACATATCTCATACGGCACAAGCTGATCTGCATAGCCTGAAAACTAGC

AACTGCCTGCTCCCTGAGCTTCCTACCAAGAAGACTCCATGTTTCTCTG

AGGAACTAGACCTTCCACCTGGACCCAGGGCCCTGAAATCCATGTCTGC

TACAGCTCAAGTCCAGGAAGTAGCCTTGGGTCAATGGTGTGTCTCCAAA

GAAAAGGAATTTCAAGAAGAAGAAAGCACAGAAGTCCCRATGCCTTTGT

ACAGTCTAAGCTTGGAAGAAGAAGAAGTGGAGGCACCGGTCTTAAAACT

CACATCTGGAGACTCTGGCTTTCATCCTGAAACCACTGACCAGGTCCTT

CAGGAGAAGAAGATGGCTCTCTTGACCTTACTCTGCTCTGCTCTGGCCT

CAAATGTGAATGTGAAAGATGCATCTGACCTTACCCGGGCATCCATCCT

TGAAGTCTGTAGTGCCCTGGCCTCCTTGGAACCGGAGTTCATCCTTAAG

GCATCTTTGTATGCTCGGCAGCAACTTAACCTCCGGGACATCGCCAATA

FIG.2B

CAGTTCTGGCTGTGGCTGCCCTCTTGCCAGCCTGCCGCCCCCATGTACG

ACGGTATTACTCCGCCATTGTTCACCTGCCTTCAGACTGGATCCAGGTA

GCCGAGTTCTACCAGAGCCTGGCAGAAGGGGATGAGAAGAAGTTGGTGT

CCCTGCCTGCCTGTCTCCGAGCTGCCATGACCGACAAATTTGCCGAGTT

TGATGAGTACCAGCTAGCTAAGTACAACCCACGGAAACATCGGTCCAAG

AGGCGGTCCCGCCAGCCACCCCGCCCTCAAAAGACAGAACGTCCATTTT

CAGAGAGAGGGAAATGTTTTCCAAAGAGCCTTTGGCCCCTTAAAAATGA

ACAGATTACGTTTGAAGCAGCTTATAATGCAATGCCAGAGAAAAACAGG

CTACCACGGTTCACTCTGAAGAAGTTGGTAGAGTATCTACATATCCACA

AGCCTGCTCAGCACGTCCAGGCCCTGCTGGGCTACAGGTACCCAGCCAC

CCTAGAGCTCTTTTCTCGGAGTCACCTCCCTGGGCCGTGGGAGTCTAGC

AGAGCTGGTCAGCGGATGAAGCTCCGAAGGCCAGAGACCTGGGAGCGGG

AGCTGAGTTTACGGGGAAACAAAGCTTCTGTGTGGGAGGAGCTCATAGA

CAATGGGAAACTGCCCTTCATGGCCATGCTCCGGAACCTGTGTAACCTG

CTGCGGACTGGGATCAGTGCCCGCCACCATGAACTCGTTCTCCAGAGAC

TCCAGCATGAGAAATCTGTGGTTCACAGTCGGCAGTTTCCATTCAGATT

CCTTAATGCTCATGACTCTATCGATAAACTTGAGGCTCAGCTCAGAAGC

AAAGCATCACCCTTCCCTTCCAATACAACATTGATGAAACGGATAATGA

TTAGAAACTCAAAAAAAAATAGGAGGCCTGCCAGTCGGAAGCACCTGTG

FIG.2C

CACCCTGACGCGCCGGCAGCTTCGGGCAGCAATGACTATACCTGTGATG

TATGAGCAGCTCAAGCGGGAGAAACTGAGGCTGCACAAGGCCAGACAAT

GGAACTGTGATGTTGAGTTGCTGGAGCGCTATCGCCAGGCCCTGGAAAC

AGCTGTGAACCTCTCAGTAAAGCACAACCTATCCCCGATGCCTGGCCGA

ACCCTCTTGGTCTATCTCACAGATGCAAATGCCGACAGGCTCTGTCCCA

AGAGTCACTCACAAGGGCCTCCCCTGAACTATGTGCTGCTGCTGATCGG

AATGATGGTGGCTCGAGCCGAGCAAGTGACTGTTTGCTTGTGTGGGGGA

GGATTTGTGAAGACACCGGTACTTACAGCCGATGAAGGCATCCTGAAGA

CTGCCATCAAACTTCAGGCTCAAGTCCAGGAGTTAGAAGGCAATGATGA

GTGGCCCCTGGACACTTTTGGGAAGTATCTGCTGTCTCTGGCTGTCCAA

AGGACCCCCATTGACAGGGTCATCCTGTTTGGTCAAAGGATGGATACCG

AGCTCCTGAAAGTAGCCAAACAGATTATCTGGCAGCATGTGAATTCCAA

GTGCCTCTTTGTTGGTGTCCTCCTACAGAAAACACAGTACATATCACCA

AATTTGAATCCCAACGATGTGACGCTCTCAGGCTGCACTGACGGGATCC

TGAAATTCATTGCCGAACATGGAGCCTCTCGTCTCCTGGAACATGTGGG

ACAACTAGATAAACTATTCAAGATCCCCCCACCCCCAGGAAAGACACAG

GCACCGTCTCTCCGGCCGCTGGAGGAGAACATCCCTGGTCCCTTGGGTC

CTATTTCCCAGCATGGATGGCGCAATATCCGGCTTTTCATTTCATCCAC

TTTCCGTGACATGCATGGGGAGCGAGATTTGCTGATGAGATCTGTTCTG

FIG.2D

CCCGCACTGCAGGCCAGAGTGTTCCCCCACCGCATCAGTCTTCACGCCA

TTGACCTGCGCTGGGGTATCACAGAGGAAGAGACCCGCAGGAACAGACA

ACTGGAAGTGTGCCTTGGGGAGGTGGAGAACTCACAGCTGTTCGTGGGG

ATTCTGGGCTCCCGCTATGGCTACATTCCCCCCAGCTATGATCTTCCTG

ATCATCCCCACTTTCACTGGACCCATGAGTACCCTTCAGGGCGATCCGT

GACAGAGATGGAGGTGATGCAATTCCTGAACCGTGGCCAACGCTCGCAG

CCTTCGGCCCAAGCTCTCATCTACTTCCGAGATCCTGATTTCCTTAGCT

CTGTGCCAGATGCCTGGAAACCTGACTTTATATCTGAGTCAGAAGAAGC

TGCACATCGGGTCTCAGAGCTGAAGAGATATCTACACGAACAGAAAGAG

GTTACCTGTCGCAGCTACTCCTGTGAATGGGGAGGTGTAGCGGCTGGCC

GGCCCTATACTGGGGGCCTGGAGGAGTTTGGACAGTTGGTTCTCCAGGA

TGTGTGGAGCATGATCCAGAAGCAGCACCTGCAGCCTGGGGCCCAGTTG

GAGCAGCCAACATCCATCTCAGAAGACGATTTGATCCAGACCAGCTTTC

AGCAGCTGAAGACCCCAACGAGTCCGGCACGGCCACGCCTTCTTCAGGA

TACAGTGCAGCAGCTGTTGCTGCCCCATGGGAGGCTGAGCCTAGTGACT

GGGCAGGCAGGACAGGGAAAGACTGCCTTTCTGGCATCCCTTGTGTCTG

CCCTGAAGGTCCCTGACCAGCCCAATGAGCCCCCGTTCGTTTTCTTCCA

CTTTGCAGCAGCCCGCCCTGACCAGTGTCTTGCTCTCAACCTCCTCAGA

CGCCTCTGTACCCATCTGCGTCAAAAACTGGGAGAGCTGAGTGCCCTCC

FIG.2E

CCAGCACTTACAGAGGCCTGGTGTGGGAACTGCAGCAGAAGTTGCTCCT

CAAATTCGCTCAGTCGCTGCAGCCTGCTCAGACTTTGGTCCTTATCATC

GATGGGGCAGATAAGTTGGTGGATCGTAATGGGCAGCTGATTTCAGACT

GGATCCCCAAGTCTCTTCCGCGGCGAGTACACCTGGTGCTGAGTGTGTC

CAGTGACTCAGGCCTGGGTGAGACCCTTCAGCAAAGTCAGGGTGCTTAT

GTGGTGGCCTTGGGCTCTTTGGTCCCATCTTCAAGGGCTCAGCTTGTGA

GAGAAGAGCTAGCACTGTATGGGAAACGACTGGAGGAGTCACCTTTTAA

CAACCAGATGCGGCTGCTGCTGGCAAAGCAGGGTTCAAGCCTGCCATTG

TACCTGCACCTTGTCACTGACTACCTGAGGCTCTTCACACTGTATGAAC

AGGTGTCTGAGAGACTTCGAACCCTGCCCGCCACTCTCCCACTGCTCTT

GCAGCACATCCTGAGCACCTTGGAGCAAGAACATGGCCATGATGTCCTT

CCTCAGGCTTTGACTGCCCTTGAGGTCACACGAAGTGGTCTGACTGTGG

ACCAGCTACATGCAATCCTGAGCACATGGCTGATCTTGCCCAAGGAGAC

TAAGAGCTGGGAAGAAGTGCTGGCTGCCAGTCACAGTGGAAACCCTTTC

CCCTTGTGTCCATTTGCCTACCTTGTCCAGAGTCTACGCAGTTTACTAG

GGGAGGGCCCAGTGGAGCGCCCTGGTGCCCGTCTCTGCCTCTCTGATGG

GCCCCTGAGGACAACAATTAAACGTCGCTATGGGAAAAGGCTGGGGCTA

GAGAAGACTGCGCATGTCCTCATTGCAGCTCACCTCTGGAAGACGTGTG

ATCCTGATGCCTCGGGCACCTTCCGAAGTTGCCCTCCTGAGGCTCTGAA

FIG.2F

AGATTTACCTTACCACCTGCTCCAGAGCGGGAACCATGGTCTCCTTGCC

GAGTTTCTTACCAATCTCCATGTGGTTGCTGCATATCTGGAAGTGGGTC

TAGTCCCCGACCTCTTGGAGGCTCATGTGCTCTATGCTTCTTCAAAGCC

TGAAGCCAACCAGAAGCTCCCAGCGGCAGATGTTGCTGTTTTCCATACC

TTCCTGAGACAACAGGCTTCACTCCTTACCCAGTATCCTTTGCTCCTGC

TCCAGCAGGCAGCTAGCCAGCCTGAAGAGTCACCTGTTTGCTGCCAGGC

CCCCCTGCTCACCCAGCGATGGCACGACCAGTTCACACTGAAATGGATT

AATAAACCCCAGACCCTGAAGGGTCAGCAAAGCTTGTCTCTGACAATGT

CCTCATCCCCAACTGCTGTGGCCTTCTCCCCGAATGGGCAAAGAGCAGC

TGTGGGGACCGCCAGTGGGACAATTTACCTGTTGAACTTGAAAACCTGG

CAGGAGGAGAAGGCTGTGGTGAGTGGCTGTGACGGGATTTCCTCTTTTG

CATTCCTTTCGGACACTGCCCTTTTCCTTACTACCTTCGACGGGCACCT

AGAGCTTTGGGACCTGCAACATGGTTGTTGGGTGTTTCAGACCAAGGCC

CACCAGTACCAAATCACTGGCTGCTGCCTGAGCCCAGACCGCCGCCTGC

TGGCCACTGTGTGTTTGGGAGGATACCTAAAGCTGTGGGACACAGTCCG

AGGACAGCTGGCTTTTCAGTACACCCATCCAAAGTCTCTCAACTGCGTT

GCCTTCCACCCAGAGGGGCAGGTGGTAGCCACAGGCAGCTGGGCTGGCA

GCATTACCTTCTTCCAGGCAGATGGACTCAAAGTCACCAAGGAACTAGG

GGCCCCCGGACCCTCTGTCTGTAGTTTGGCATTCAACAAACCTGGGAAG

FIG.2G

ATTGTGGCTGTGGGCCGGATAGATGGGACAGTGGAGCTGTGGGCCTGGC

AAGAGGGTGCCCGGCTGGCGGCCTTCCCTGCACAGTGTGGCTGTGTCTC

TGCTGTTCTTTTCTTGCATGCTGGAGACCGGTTCCTGACTGCTGGAGAA

GATGGCAAGGCTCAGTTATGGTCAGGATTTCTTGGCCGGCCCAGGGGTT

GCCTGGGCTCTCTTCCTCTTTCTCCTGCACTCTCGGTGGCTCTCAACCC

AGACGGTGACCAGGTGGCTGTTGGGTACCGAGAAGATGGCATTAACATC

TACAAGATTTCTTCAGGTTCCCAGGGGCCTCAGCATCAAGAGCTAAATG

TGGCGGTGTCTGCACTGGTGTGGCTGAGCCCTAGTGTTTTGGTGAGTGG

TGCAGAAGATGGATCCCTGCATGGTTGGATGTTCAAGGGAGACTCCCTT

CATTCCCTGTGGCTGTTGTCGAGATACCAGAAGCCTGTGCTGGGACTGG

CTGCCTCCCGGGAACTCATGGCTGCTGCCTCAGAGGACTTCACTGTGAG

ACTGTGGCCCAGACAGCTGCTGACACAGCCACATGTGCATGCGGTAGAG

TTGCCCTGTTGTGCTGAACTCCGGGGACACGAGGGGCCAGTGTGCTGCT

GTAGCTTCAGCCCTGATGGAGGCATCTTGGCCACAGCTGGCAGGGATCG

GAATCTCCTTTGCTGGGACATGAAGATAGCCCAAGCCCCTCTCCTGATT

CACACTTTCTCGTCCTGTCATCGTGACTGGATCACTGGCTGTGCGTGGA

CCAAAGACAACATCCTGGTCTCCTGCTCGAGTGATGGCTCTGTGGGACT

CTGGAACCCAGAGGCAGGGCAGCAACTTGGCCAGTTCTCAGGCCACCAG

AGTGCCGTGAGCGCCGTGGTTGCTGTGGAGGAACACATTGTATCTGTGA

FIG.2H

GCCGAGATGGGACCTTGAAAGTGTGGGACCATCAGGGTGTGGAGCTGAC

CAGCATCCCTGCCCATTCCGGACCCATCAGCCAGTGTGCAGCTGCTCTG

GAGCCCCGCCCAGGGGGACAGCCTGGATCAGAGCTTCTGGTGGTGACTG

TTGGACTAGATGGGGCCACAAAGTTGTGGCATCCCCTGTTGGTGTGCCA

AATACGTACTCTCCAGGGACACAGTGGCCCAGTCACAGCAGCTGCTGCT

TCAGAGGCCTCAGGCCTCCTGCTGACCTCAGATGATAGCTCTGTACAGC

TCTGGCAGATACCAAAGGAAGCAGATGATTCATACAAACCTAGGAGTTC

TGTGGCCATCACTGCTGTGGCATGGGCACCGGATGGTTCTATGGTGGTG

TCCGGAAATGAAGCCGGGGAACTGACACTGTGGCAGCAAGCCAAGGCTG

TGGCTACCGCACAGGCTCCAGGCCGCGTCAGTCACCTGATCTGGTACTC

GGCAAATTCATTCTTCGTTCTCAGTGCTAATGAAAACGTCAGCGAGTGG

CAAGTGGGACTGAGGAAAGGTTCAACGTCCACCAGTTCCAGTCTTCATC

TGAAGAGAGTTCTGCAGGAGGACTGGGGAGTCTTGACAGGTCTGGGTCT

GGCCCCTGATGGCCAGTCTCTCATCTTGATGAAAGAGGATGTGGAATTA

CTAGAGATGAAGCCTGGGTCTATTCCATCTTCTATCTGCAGGAGGTATG

GAGTACATTCTTCAATACTGTGCACCAGCAAGGAGTACGGCTTGTTCTA

CCTGCAGCAGGGGGACTCCGGATTACTTTCTATATTGGAGCAAAAGGAG

TCAGGGGAGTTTGAAGAGATCCTGGACTTCAATCTGAACTTAAATAATC

CTAATGGGTCCCCAGTATCAATCACTCAGGCCAAACCTGAGTCTGAATC

FIG.2I

ATCCCTTTTGTGCGCCACCTCTGATGGGATGCTGTGGAACTTATCTGAA

TGTACCTCAGAGGGAGAATGGATCGTAGATAACATTTGGCAGAAAAAAG

CAAAAAAACCTAAAACTCAGACTCTGGAGACAGAGTTGTCCCCGCACTC

AGAGTTGGATTTTTCCATTGATTGCTGGATTGATCCCACAAATTTAAAG

GCACAGCAGTGTAAAAAGATCCACTTGGGCTCTGTCACAGCCCTCCATG

TGCTTCCGGGATTGCTGGTGACAGCTTCGAAGGACAGAGATGTTAAGCT

GTGGGAGAGACCCAGTATGCAGCTGCTGGGCTTGTTCCGATGTGAAGGG

CCAGTGAGCTGTCTGGAACCTTGGATGGAGCCCAGCTCTCCCCTGCAGC

TTGCTGTGGGAGACACACAAGGAAACTTGTATTTTCTATCTTGGGAA

FIG.3A

MEKLHGHVSAHPDILSLENRCLAMLPDLQPLEKLHQHVSTHSDILSLKN

QCLATLPDLKTMEKPHGYVSAHPDILSLENQCLATLSDLKTMEKPHGHV

SAHPDILSLENRCLATLPSLKSTVSASPLFQSLQISHMTQADLYRVNNS

NCLLSEPPSWRAQHFSKGLDLSTCPIALKSISATETAQEATLGRWFDSE

EKKGAETQMPSYSLSLGEEEEVEDLAVKLTSGDSESHPEPTDHVLQEKK

MALLSLLCSTLVSEVNMNNTSDPTLAAIFEICRELALLEPEFILKASLY

ARQQLNVRNVANNILAIAAFLPACRPHLRRYFCAIVQLPSDWIQVAELY

QSLAEGDKNKLVPLPACLRTAMTDKFAQFDEYQLAKYNPRKHRAKRHPR

RPPRSPGMEPPFSHRCFPRYIGFLREEQRKFEKAGDTVSEKKNPPRFTL

KKLVQRLHIHKPAQHVQALLGYRYPSNLQLFSRSRLPGPWDSSRAGKRM

KLSRPETWERELSLRGNKASVWEELIENGKLPFMAMLRNLCNLLRVGIS

SRHHELILQRLQHGKSVIHSRQFPFRFLNAHDAIDALEAQLRNQALPFP

SNITLMRRILTRNEKNRPRRRFLCHLSRQQLRMAMRIPVLYEQLKREKL

RVHKARQWKYDGEMLNRYRQALETAVNLSVKHSLPLLPGRTVLVYLTDA

NADRLCPKSNPQGPPLNYALLLIGMMITRAEQVDVVLCGGDTLKTAVLK

AEEGILKTAIKLQAQVQEFDENDGWSLNTFGKYLLSLAGQRVPVDRVIL

LGQSMDDGMINVAKQLYWQRVNSKCLFVGILLRRVQYLSTDLNPNDVTL

SGCTDAILKFIAEHGASHLLEHVGQMDKIFKIPPPPGKTGVQSLRPLEE

DTPSPLAPVSQQGWRSIRLFISSTFRDMHGERDLLLRSVLPALQARAAP

FIG.3B

HRISLHGIDLRWGVTEEETRRNRQLEVCLGEVENAQLFVGILGSRYGYI

PPSYNLPDHPHFHWAQQYPSGRSVTEMEVMQFLNRNQRLQPSAQALIYF

RDSSFLSSVPDAWKSDFVSESEEAAXRISELKSYLSRQKGITCRRYPCE

WGGVAAGRPYVGGLEEFGQLVLQDVWNMIQKLYLQPGALLEQPVSIPDD

DLVQATFQQLQKPPSPARPRLLQDTVQXLMLPHGRLSLVTGQSGQGKTA

FLASLVSALQAPDGAKVAXLVFFHFSGARPDQGLALTLLRRLCTYLRGQ

LKEPGALPSTYRSLVWELQQRLLPKSAESLHPGQTQVLIIDGADRLVDQ

NGQLISDWIPKKLPRCVHLVLSVSSDAGLGETLEQSQGAHVLALGPLEA

SARARLVREELALYGKRLEESPFNNQMRLLLVKRESGRPLYLRLVTDHL

RLFTLYEQVSERLRTLPATVPLLLQHILSTLEKEHGPDVLPQALTALEV

TRSGLTVDQLHGVLSVWRTLPKGTKSWEEAVAAGNSGDPYPMGPFACLV

QSLRSLLGEGPLERPGARLCLPDGPLRTAAKRCYGKRPGLEDTAHILIA

AQLWKTCDADASGTFRSCPPEALGDLPYHLLQSGNRGLLSKFLTNLHVV

AAHLELGLVSRLLEAHALYASSVPKEEQKLPEADVAVFRTFLRQQASIL

SQYPRLLPQQAANQPLDSPLCHQASLLSRRWHLQHTLRWLNKPRTMKNQ

QSSSLSLAVSSSPTAVAFSTNGQRAAVGTANGTVYLLDLRTWQEEKSVV

SGCDGISACLFLSDDTLFLTAFDGLLELWDLQHGCRVLQTKAHQYQITG

CCLSPDCRLLATVCLGGCLKLWDTVRGQLAFQHTYPKSLNCVAFHPEGQ

VIATGSWAGSISFFQVDGLKVTKDLGAPGASIRTLAFNVPGGVVAVGRL

FIG.3C

DSMVELWAWREGARLAAFPAHHGFVAAALFLHAGCQLLTAGEDGKVQVW

SGSLGRPRGHLGSLSLSPALSVALSPDGDRVAVGYRADGIRIYKISSGS

QGAQGQALDVAVSALAWLSPKVLVSGAEDGSLQGWALKECSLQSLWLLS

RFQKPVLGLATSQELLASASEDFTVQLWPRQLLTRPHKAEDFPCGTELR

GHEGPVSCCSFSTDGGSLATGGRDRSLLCWDVRTPKTPVLIHSFPACHR

DWVTGCAWTKDNLLISCSSDGSVGLWDPESGQRLGQFLGHQSAVSAVAA

VEEHVVSVSRDGTLKVWDHQGVELTSIPAHSGPISHCAAAMEPRAAGQP

GSELLVVTVGLDGATRLWHPLLVCQTHTLLGHSGPVRAAAVSETSGLML

TASEDGSVRLWQVPKEADDTCIPRSSAAVTAVAWAPDGSMAVSGNQAGE

LILWQEAKAVATAQAPGHIGALIWSSAHTFFVLSADEKISEWQVKLRKG

SAPGNLSLHLNRILQEDLGVLTSLDWAPDGHFLILAKADLKLLCMKPGD

APSEIWSSYTENPMILSTHKEYGIFVLQPKDPGVLSFLRQKESGEFEER

LNFDINLENPSRTLISITQAKPESESSFLCASSDGILWNLAKCSPEGEW

TTGNMWQKKANTPETQTPGTDPSTCRESDASMDSDASMDSEPTPHLKTR

QRRKIHSGSVTALHVLPELLVTASKDRDVKLWERPSMQLLGLFRCEGSV

SCLEPWLGANSTLQLAVGDVQGNVYFLNWE

FIG.4A

MEKLCGHVPGHSDILSLKNRCLTMLPDLQPLEKIHGHRSVHSDILSLEN

QCLTMLSDLQPTERIDGHISVHPDILSLENRCLTMLPDLQPLEKLCGHM

SSHPDVLSLENQCLATLPTVKSTALTSPLLQGLHISHTAQADLHSLKTS

NCLLPELPTKKTPCFSEELDLPPGPRALKSMSATAQVQEVALGQWCVSK

EKEFQEEESTEVPMPLYSLSLEEEEVEAPVLKLTSGDSGFHPETTDQVL

QEKKMALLTLLCSALASNVNVKDASDLTRASILEVCSALASLEPEFILK

ASLYARQQLNLRDIANTVLAVAALLPACRPHVRRYYSAIVHLPSDWIQV

AEFYQSLAEGDEKKLVSLPACLRAAMTDKFAEFDEYQLAKYNPRKHRSK

RRSRQPPRPQKTERPFSERGKCFPKSLWPLKNEQITFEAAYNAMPEKNR

LPRFTLKKLVEYLHIHKPAQHVQALLGYRYPATLELFSRSHLPGPWESS

RAGQRMKLRRPETWERELSLRGNKASVWEELIDNGKLPFMAMLRNLCNL

LRTGISARHHELVLQRLQHEKSVVHSRQFPFRFLNAHDSIDKLEAQLRS

KASPFPSNTTLMKRIMIRNSKKNRRPASRKHLCTLTRRQLRAAMTIPVM

YEQLKREKLRLHKARQWNCDVELLERYRQALETAVNLSVKHNLSPMPGR

TLLVYLTDANADRLCPKSHSQGPPLNYVLLLIGMMVARAEQVTVCLCGG

GFVKTPVLTADEGILKTAIKLQAQVQELEGNDEWPLDTFGKYLLSLAVQ

RTPIDRVILFGQRMDTELLKVAKQIIWQHVNSKCLFVGVLLQKTQYISP

NLNPNDVTLSGCTDGILKFIAEHGASRLLEHVGQLDKLFKIPPPPGKTQ

APSLRPLEENIPGPLGPISQHGWRNIRLFISSTFRDMHGERDLLMRSVL

FIG.4B

PALQARVFPHRISLHAIDLRWGITEEETRRNRQLEVCLGEVENSQLFVG

ILGSRYGYIPPSYDLPDHPHFHWTHEYPSGRSVTEMEVMQFLNRGQRSQ

PSAQALIYFRDPDFLSSVPDAWKPDFISESEEAAHRVSELKRYLHEQKE

VTCRSYSCEWGGVAAGRPYTGGLEEFGQLVLQDVWSMIQKQHLQPGAQL

EQPTSISEDDLIQTSFQQLKTPTSPARPRLLQDTVQQLLLPHGRLSLVT

GQAGQGKTAFLASLVSALKVPDQPNEPPFVFFHFAAARPDQCLALNLLR

RLCTHLRQKLGELSALPSTYRGLVWELQQKLLLKFAQSLQPAQTLVLII

DGADKLVDRNGQLISDWIPKSLPRRVHLVLSVSSDSGLGETLQQSQGAY

VVALGSLVPSSRAQLVREELALYGKRLEESPFNNQMRLLLAKQGSSLPL

YLHLVTDYLRLFTLYEQVSERLRTLPATLPLLLQHILSTLEQEHGHDVL

PQALTALEVTRSGLTVDQLHAILSTWLILPKETKSWEEVLAASHSGNPF

PLCPFAYLVQSLRSLLGEGPVERPGARLCLSDGPLRTTIKRRYGKRLGL

EKTAHVLIAAHLWKTCDPDASGTFRSCPPEALKDLPYHLLQSGNHGLLA

EFLTNLHVVAAYLEVGLVPDLLEAHVLYASSKPEANQKLPAADVAVFHT

FLRQQASLLTQYPLLLLQQAASQPEESPVCCQAPLLTQRWHDQFTLKWI

NKPQTLKGQQSLSLTMSSSPTAVAFSPNGQRAAVGTASGTIYLLNLKTW

QEEKAVVSGCDGISSFAFLSDTALFLTTFDGHLELWDLQHGCWVFQTKA

HQYQITGCCLSPDRRLLATVCLGGYLKLWDTVRGQLAFQYTHPKSLNCV

AFHPEGQVVATGSWAGSITFFQADGLKVTKELGAPGPSVCSLAFNKPGK

FIG.4C

IVAVGRIDGTVELWAWQEGARLAAFPAQCGCVSAVLFLHAGDRFLTAGE

DGKAQLWSGFLGRPRGCLGSLPLSPALSVALNPDGDQVAVGYREDGINI

YKISSGSQGPQHQELNVAVSALVWLSPSVLVSGAEDGSLHGWMFKGDSL

HSLWLLSRYQKPVLGLAASRELMAAASEDFTVRLWPRQLLTQPHVHAVE

LPCCAELRGHEGPVCCCSFSPDGGILATAGRDRNLLCWDMKIAQAPLLI

HTFSSCHRDWITGCAWTKDNILVSCSSDGSVGLWNPEAGQQLGQFSGHQ

SAVSAVVAVEEHIVSVSRDGTLKVWDHQGVELTSIPAHSGPISQCAAAL

EPRPGGQPGSELLVVTVGLDGATKLWHPLLVCQIRTLQGHSGPVTAAAA

SEASGLLLTSDDSSVQLWQIPKEADDSYKPRSSVAITAVAWAPDGSMVV

SGNEAGELTLWQQAKAVATAQAPGRVSHLIWYSANSFFVLSANENVSEW

QVGLRKGSTSTSSSLHLKRVLQEDWGVLTGLGLAPDGQSLILMKEDVEL

LEMKPGSIPSSICRRYGVHSSILCTSKEYGLFYLQQGDSGLLSILEQKE

SGEFEEILDFNLNLNNPNGSPVSITQAKPESESSLLCATSDGMLWNLSE

CTSEGEWIVDNIWQKKAKKPKTQTLETELSPHSELDFSIDCWIDPTNLK

AQQCKKIHLGSVTALHVLPGLLVTASKDRDVKLWERPSMQLLGLFRCEG

PVSCLEPWMEPSSPLQLAVGDTQGNLYFLSWE

FIG.5A

CACGCGTCCGGGCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGG

CCCCGGCCACCCCCGCGATGCCGCGCGCTCCCCGCTGCCGAGCCGTGCG

CTCCCTGCTGCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCCACGTTC

GTGCGGCGCCTGGGGCCCCAGGGCTGGCGGCTGGTGCAGCGCGGGGACC

CGGCGGCTTTCCGCGCGCTGGTGGCCCAGTGCCTGGTGTGCGTGCCCTG

GGACGCACGGCCGCCCCCCGCCGCCCCCTCCTTCCGCCAGGTGTCCTGC

CTGAAGGAGCTGGTGGCCCGAGTGCTGCAGAGGCTGTGCGAGCGCGGCG

CGAAGAACGTGCTGGCCTTCGGCTTCGCGCTGCTGGACGGGGCCCGCGG

GGGCCCCCCCCGAGGCCTTCACCACCAGCGTGCGCAGCTACCTGCCCAAC

ACGGTGACCGACGCACTGCGGGGGAGCGGGGCGTGGGGGCTGCTGCTGC

GCCGCGTGGGCGACGACGTGCTGGTTCACCTGCTGGCACGCTGCGCGCT

CTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCCG

CTGTACCAGCTCGGCGCTGCCACTCAGGCCCGGCCCCCGCCACACGCTA

GTGGACCCCGAAGGCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGT

CAGGGAGGCCGGGGTCCCCCTGGGCCTGCCAGCCCCGGGTGCGAGGAGG

CGCGGGGGCAGTGCCAGCCGAAGTCTGCCGTTGCCCAAGAGGCCCAGGC

GTGGCGCTGCCCCTGAGCCGGAGCGGACGCCCGTTGGGCAGGGGTCCTG

GGCCCACCCGGGCAGGACGCGTGGACCGAGTGACCGTGGTTTCTGTGTG

GTGTCACCTGCCAGACCCGCCGAAGAAGCCACCTCTTTGGAGGGTGCGC

FIG.5B

TCTCTGGCACGCGCCACTCCCACCCATCCGTGGGCCGCCAGCACCACGC

GGGCCCCCCATCCACATCGCGGCCACCACGTCCCTGGGACACGCCTTGT

CCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTCCTCAGGCGACA

AGGAGCAGCTGCGGCCCTCCTTCCTACTCAGCTCTCTGAGGCCCAGCCT

GACTGGCGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCCAGGCCC

TGGATGCCAGGGACTCCCCGCAGGTTGCCCCGCCTGCCCCAGCGCTACT

GGCAAATGCGGCCCCTGTTTCTGGAGCTGCTTGGGAACCACGCGCAGTG

CCCCTACGGGGTGCTCCTCAAGACGCACTGCCCGCTGCGAGCTGCGGTC

ACCCCAGCAGCCGGTGTCTGTGCCCGGGAGAAGCCCCAGGGCTCTGTGG

CGGCCCCCGAGGAGGAGGACACAGACCCCCGTCGCCTGGTGCAGCTGCT

CCGCCAGCACAGCAGCCCCTGGCAGGTGTACGGCTTCGTGCGGGCCTGC

CTGCGCCGGCTGGTGCCCCCAGGCCTCTGGGGCTCCAGGCACAACGAAC

GCCGCTTCCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAGCATGC

CAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCGGGACTGC

GCTTGGCTGCGCAGGAGCCCAGGGGTTGGCTGTGTTCCGGCCGCAGAGC

ACCGTCTGCGTGAGGAGATCCTGGCCAAGTTCCTGCACTGGCTGATGAG

TGTGTACGTCGTCGAGCTGCTCAGGTCTTTCTTTTATGTCACGGAGACC

ACGTTTCAAAAGAACAGGCTCTTTTTCTACCGGAAGAGTGTCTGGAGCA

AGTTGCAAAGCATTGGAATCAGACAGCACTTGAAGAGGGTGCAGCTGCG

FIG.5C

GGAGCTGTCGGAAGCAGAGGTCAGGCAGCATCGGGAAGCCAGGCCCGCC

CTGCTGACGTCCAGACTCCGCTTCATCCCCAAGCCTGACGGGCTGCGGC

CGATTGTGAACATGGACTACGTCGTGGGAGCCAGAACGTTCCGCAGAGA

AAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTTCAGCGTG

CTCAACTACGAGCGGGCGCGGCGCCCCGGCCTCCTGGGCGCCTCTGTGC

TGGGCCTGGACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGT

GCGGGCCCAGGACCCGCCGCCTGAGCTGTACTTTGTCAAGGTGGATGTG

ACGGGCGCGTACGACACCATCCCCCAGGACAGGCTCACGGAGGTCATCG

CCAGCATCATCAAACCCCAGAACACGTACTGCGTGCGTCGGTATGCCGT

GGTCCAGAAGGCCGCCCATGGGCACGTCCGCAAGGCCTTCAAGAGCCAC

GTCTCTACCTTGACAGACCTCCAGCCGTACATGCGACAGTTCGTGGCTC

ACCTGCAGGAGACCAGCCCGCTGAGGGATGCCGTCGTCATCGAGCAGAG

CTCCTCCCTGAATGAGGCCAGCAGTGGCCTCTTCGACGTCTTCCTACGC

TTCATGTGCCACCACGCCGTGCGCATCAGGGGCAAGTCCTACGTCCAGT

GCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCTGCAGCCT

GTGCTACGGCGACATGGAGAACAAGCTGTTTGCGGGGATTCGGCGGGAC

GGGCTGCTCCTGCGTTTGGTGGATGATTTCTTGTTGGTGACACCTCACC

TCACCCACGCGAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGA

GTATGGCTGCGTGGTGAACTTGCGGAAGACAGTGGTGAACTTCCCTGTA

FIG.5D

GAAGACGAGGCCCTGGGTGGCACGGCTTTTGTTCAGATGCCGGCCCACG

GCCTAT

FIG.6A

HASGQRCVLLRTWEALAPATPAMPRAPRCRAVRSLLRSHYREVLPLATF

VRRLGPQGWRLVQRGDPAAFRALVAQCLVCVPWDARPPPAAPSFRQVSC

LKELVARVLQRLCERGAKNVLAFGFALLDGARGGPPEAFTTSVRSYLPN

TVTDALRGSGAWGLLLRRVGDDVLVHLLARCALFVLVAPSCAYQVCGPP

LYQLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPGARR

RGGSASRSLPLPKRPRRGAAPEPERTPVGQGSWAHPGRTRGPSDRGFCV

VSPARPAEEATSLEGALSGTRHSHPSVGRQHHAGPPSTSRPPRPWDTPC

PPVYAETKHFLYSSGDKEQLRPSFLLSSLRPSLTGARRLVETIFLGSRP

WMPGTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPLRAAV

TPAAGVCAREKPQGSVAAPEEEDTDPRRLVQLLRQHSSPWQVYGFVRAC

LRRLVPPGLWGSRHNERRFLRNTKKFISLGKHAKLSLQELTWKMSVRDC

AWLRRSPGVGCVPAAEHRLREEILAKFLHWLMSVYVVELLRSFFYVTET

TFQKNRLFFYRKSVWSKLQSIGIRQHLKRVQLRELSEAEVRQHREARPA

LLTSRLRFIPKPDGLRPIVNMDYVVGARTFRREKRAERLTSRVKALFSV

LNYERARRPGLLGASVLGLDDIHRAWRTFVLRVRAQDPPPELYFVKVDV

TGAYDTIPQDRLTEVIASIIKPQNTYCVRRYAVVQKAAHGHVRKAFKSH

VSTLTDLQPYMRQFVAHLQETSPLRDAVVIEQSSSLNEASSGLFDVFLR

FMCHHAVRIRGKSYVQCQGIPQGSILSTLLCSLCYGDMENKLFAGIRRD

FIG.6B

GLLLRLVDDFLLVTPHLTHAKTFLRTLVRGVPEYGCVVNLRKTVVNFPV

EDEALGGTAFVQMPAHGL

FIG. 7

TCCCCTGGTGCGGCCTGCTGCTGGATACCCGGACCCTGGAGGTGCAGAGCGACT

ACTCCAGCTATGCCCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCGCGGCT

TCAAGGCTGGGAGGAACATGCGTCGCAAACTCTTTGGGGTCTTGCGGCTGAAGT

GTCACAGCCTGTTTCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCA

ACATCTACAAGATCCTCCTGCTGCAGGCGTACAGGTTTCACGCATGTGTGCTGC

AGCTCCCATTTCATCAGCAAGTTTGGAAGAACCCCACATTTTTCCTGCGCGTCA

TCTCTGACACGGCCTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGGGA

TGTCGCTGGGGGCCAAGGGCGCCGCCGGCCCTCTGCCCTCCGAGGCCGTGCAGT

GGCTGTGCCACCAAGCATTCCTGCTCAAGCTGACTCGACACCGTGTCACCTACG

TGCCACTCCTGGGGTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCC

CGGGGACGACGCTGACTGCCCTGGAGGCCGCAGCCAACCCGGCACTGCCCTCAG

ACTTCAAGACCATCCTGGACTGATGGCCACCCGCCCACAGCCAGGCCGAGAGCA

GACACCAGCAGCCCTGTCACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGC

CCACACCCAGGCCCGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTTTGGCCG

AGGCCTGCATGTCCGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGT

CCAGCCAAGGGCTGAGTGTCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCT

GGCGCTCGGCTCCACCCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCA

CTCCCCACATAGGAATAGTCCATCCCCTGAT

FIG.8A

CCACGCGTCCGGGCAGCGCTGCGTCCTGCTGCGCACGTGGGAAGCCCTGGCCCC

GGCCACCCCCGCGATGCCGCGCGCTCCCCGCTGCCGAGCCGTGCGCTCCCTGCT

GCGCAGCCACTACCGCGAGGTGCTGCCGCTGGCCACGTTCGTGCGGCGCCTGGG

GCCCCAGGGCTGGCGGCTGGTGCAGCGCGGGGACCCGGCGGCTTTCCGCGCGCT

GGTGGCCCAGTGCCTGGTGTGCGTGCCCTGGGACGCACGGCCGCCCCCCGCCGC

CCCCTCCTTCCGCCAGGTGTCCTGCCTGAAGGAGCTGGTGGCCCGAGTGCTGCA

GAGGCTGTGCGAGCGCGGCGCGAAGAACGTGCTGGCCTTCGGCTTCGCGCTGCT

GGACGGGGCCCGCGGGGGCCCCCCCCGAGGCCTTCACCACCAGCGTGCGCAGCTA

CCTGCCCAACACGGTGACCGACGCACTGCGGGGGAGCGGGGCGTGGGGGCTGCT

GCTGCGCCGCGTGGGCGACGACGTGCTGGTTCACCTGCTGGCACGCTGCGCGCT

CTTTGTGCTGGTGGCTCCCAGCTGCGCCTACCAGGTGTGCGGGCCGCCGCTGTA

CCAGCTCGGCGCTGCCACTCAGGCCCGGCCCCCGCCACACGCTAGTGGACCCCG

AAGGCGTCTGGGATGCGAACGGGCCTGGAACCATAGCGTCAGGGAGGCCGGGGT

CCCCCTGGGCCTGCCAGCCCCGGGTGCGAGGAGGCGCGGGGGCAGTGCCAGCCG

AAGTCTGCCGTTGCCCAAGAGGCCCAGGCGTGGCGCTGCCCCTGAGCCGGAGCG

GACGCCCGTTGGGCAGGGGTCCTGGGCCCACCCGGGCAGGACGCGTGGACCGAG

TGACCGTGGTTTCTGTGTGGTGTCACCTGCCAGACCCGCCGAAGAAGCCACCTC

TTTGGAGGGTGCGCTCTCTGGCACGCGCCACTCCCACCCATCCGTGGGCCGCCA

GCACCACGCGGGCCCCCCATCCACATCGCGGCCACCACGTCCCTGGGACACGCC

TTGTCCCCCGGTGTACGCCGAGACCAAGCACTTCCTCTACTCCTCAGGCGACAA

FIG.8B

GGAGCAGCTGCGGCCCTCCTTCCTACTCAGCTCTCTGAGGCCCAGCCTGACTGG

CGCTCGGAGGCTCGTGGAGACCATCTTTCTGGGTTCCAGGCCCTGGATGCCAGG

GACTCCCCGCAGGTTGCCCCGCCTGCCCCAGCGCTACTGGCAAATGCGGCCCCT

GTTTCTGGAGCTGCTTGGGAACCACGCGCAGTGCCCCTACGGGGTGCTCCTCAA

GACGCACTGCCCGCTGCGAGCTGCGGTCACCCCAGCAGCCGGTGTCTGTGCCCG

GGAGAAGCCCCAGGGCTCTGTGGCGGCCCCCGAGGAGGAGGACACAGACCCCCG

TCGCCTGGTGCAGCTGCTCCGCCAGCACAGCAGCCCCTGGCAGGTGTACGGCTT

CGTGCGGGCCTGCCTGCGCCGGCTGGTGCCCCCAGGCCTCTGGGGCTCCAGGCA

CAACGAACGCCGCTTCCTCAGGAACACCAAGAAGTTCATCTCCCTGGGGAAGCA

TGCCAAGCTCTCGCTGCAGGAGCTGACGTGGAAGATGAGCGTGCGGGACTGCGC

TTGGCTGCGCAGGAGCCCAGGGGTTGGCTGTGTTCCGGCCGCAGAGCACCGTCT

GCGTGAGGAGATCCTGGCCAAGTTCCTGCACTGGCTGATGAGTGTGTACGTCGT

CGAGCTGCTCAGGTCTTTCTTTTATGTCACGGAGACCACGTTTCAAAAGAACAG

GCTCTTTTTTCTACCGGAAGAGTGTCTGGAGCAAGTTGCAAAGCATTGGAATCAG

ACAGCACTTGAAGAGGGTGCAGCTGCGGGAGCTGTCGGAAGCAGAGGTCAGGCA

GCATCGGGAAGCCAGGCCCGCCCTGCTGACGTCCAGACTCCGCTTCATCCCCAA

GCCTGACGGGCTGCGGCCGATTGTGAACATGGACTACGTCGTGGGAGCCAGAAC

GTTCCGCAGAGAAAAGAGGGCCGAGCGTCTCACCTCGAGGGTGAAGGCACTGTT

CAGCGTGCTCAACTACGAGCGGGCGCGGCGCCCCGGCCTCCTGGGCGCCTCTGT

GCTGGGCCTGGACGATATCCACAGGGCCTGGCGCACCTTCGTGCTGCGTGTGCG

FIG.8C

GGCCCAGGACCCGCCGCCTGAGCTGTACTTTGTCAAGGTGGATGTGACGGGCGC

GTACGACACCATCCCCCAGGACAGGCTCACGGAGGTCATCGCCAGCATCATCAA

ACCCCAGAACACGTACTGCGTGCGTCGGTATGCCGTGGTCCAGAAGGCCGCCCA

TGGGCACGTCCGCAAGGCCTTCAAGAGCCACGTCTCTACCTTGACAGACCTCCA

GCCGTACATGCGACAGTTCGTGGCTCACCTGCAGGAGACCAGCCCGCTGAGGGA

TGCCGTCGTCATCGAGCAGAGCTCCTCCCTGAATGAGGCCAGCAGTGGCCTCTT

CGACGTCTTCCTACGCTTCATGTGCCACCACGCCGTGCGCATCAGGGGCAAGTC

CTACGTCCAGTGCCAGGGGATCCCGCAGGGCTCCATCCTCTCCACGCTGCTCTG

CAGCCTGTGCTACGGCGACATGGAGAACAAGCTGTTTGCGGGGATTCGGCGGGA

CGGGCTGCTCCTGCGTTTGGTGGATGATTTCTTGTTGGTGACACCTCACCTCAC

CCACGCGAAAACCTTCCTCAGGACCCTGGTCCGAGGTGTCCCTGAGTATGGCTG

CGTGGTGAACTTGCGGAAGACAGTGGTGAACTTCCCTGTAGAAGACGAGGCCCT

GGGTGGCACGGCTTTTGTTCAGATGCCGGCCCACGGCCTATTCCCCTGGTGCGG

CCTGCTGCTGGATACCCGGACCCTGGAGGTGCAGAGCGACTACTCCAGCTATGC

CCGGACCTCCATCAGAGCCAGTCTCACCTTCAACCGCGGCTTCAAGGCTGGGAG

GAACATGCGTCGCAAACTCTTTGGGGTCTTGCGGCTGAAGTGTCACAGCCTGTT

TCTGGATTTGCAGGTGAACAGCCTCCAGACGGTGTGCACCAACATCTACAAGAT

CCTCCTGCTGCAGGCGTACAGGTTTCACGCATGTGTGCTGCAGCTCCCATTTCA

TCAGCAAGTTTGGAAGAACCCCACATTTTTCCTGCGCGTCATCTCTGACACGGC

CTCCCTCTGCTACTCCATCCTGAAAGCCAAGAACGCAGGGATGTCGCTGGGGGC

FIG.8D

CAAGGGCGCCGCCGGCCCTCTGCCCTCCGAGGCCGTGCAGTGGCTGTGCCACCA

AGCATTCCTGCTCAAGCTGACTCGACACCGTGTCACCTACGTGCCACTCCTGGG

GTCACTCAGGACAGCCCAGACGCAGCTGAGTCGGAAGCTCCCGGGGACGACGCT

GACTGCCCTGGAGGCCGCAGCCAACCCGGCACTGCCCTCAGACTTCAAGACCAT

CCTGGACTGATGGCCACCCGCCCACAGCCAGGCCGAGAGCAGACACCAGCAGCC

CTGTCACGCCGGGCTCTACGTCCCAGGGAGGGAGGGGCGGCCCACACCCAGGCC

CGCACCGCTGGGAGTCTGAGGCCTGAGTGAGTGTTTGGCCGAGGCCTGCATGTC

CGGCTGAAGGCTGAGTGTCCGGCTGAGGCCTGAGCGAGTGTCCAGCCAAGGGCT

GAGTGTCCAGCACACCTGCCGTCTTCACTTCCCCACAGGCTGGCGCTCGGCTCC

ACCCCAGGGCCAGCTTTTCCTCACCAGGAGCCCGGCTTCCACTCCCCACATAGG

AATAGTCCATCCCCTGAT

FIG.9A

HASGQRCVLLRTWEALAPATPAMPRAPRCRAVRSLLRSHYREVLPLATF

VRRLGPQGWRLVQRGDPAAFRALVAQCLVCVPWDARPPPAAPSFRQVSC

LKELVARVLQRLCERGAKNVLAFGFALLDGARGGPPEAFTTSVRSYLPN

TVTDALRGSGAWGLLLRRVGDDVLVHLLARCALFVLVAPSCAYQVCGPP

LYQLGAATQARPPPHASGPRRRLGCERAWNHSVREAGVPLGLPAPGARR

RGGSASRSLPLPKRPRRGAAPEPERTPVGQGSWAHPGRTRGPSDRGFCV

VSPARPAEEATSLEGALSGTRHSHPSVGRQHHAGPPSTSRPPRPWDTPC

PPVYAETKHFLYSSGDKEQLRPSFLLSSLRPSLTGARRLVETIFLGSRP

WMPGTPRRLPRLPQRYWQMRPLFLELLGNHAQCPYGVLLKTHCPLRAAV

TPAAGVCAREKPQGSVAAPEEEDTDPRRLVQLLRQHSSPWQVYGFVRAC

LRRLVPPGLWGSRHNERRFLRNTKKFISLGKHAKLSLQELTWKMSVRDC

AWLRRSPGVGCVPAAEHRLREEILAKFLHWLMSVYVVELLRSFFYVTET

TFQKNRLFFYRKSVWSKLQSIGIRQHLKRVQLRELSEAEVRQHREARPA

LLTSRLRFIPKPDGLRPIVNMDYVVGARTFRREKRAERLTSRVKALFSV

LNYERARRPGLLGASVLGLDDIHRAWRTFVLRVRAQDPPPELYFVKVDV

TGAYDTIPQDRLTEVIASIIKPQNTYCVRRYAVVQKAAHGHVRKAFKSH

VSTLTDLQPYMRQFVAHLQETSPLRDAVVIEQSSSLNEASSGLFDVFLR

FMCHHAVRIRGKSYVQCQGIPQGSILSTLLCSLCYGDMENKLFAGIRRD

GLLLRLVDDFLLVTPHLTHAKTFLRTLVRGVPEYGCVVNLRKTVVNFPV

FIG.9B

EDEALGGTAFVQMPAHGLFPWCGLLLDTRTLEVQSDYSSYARTSIRASL

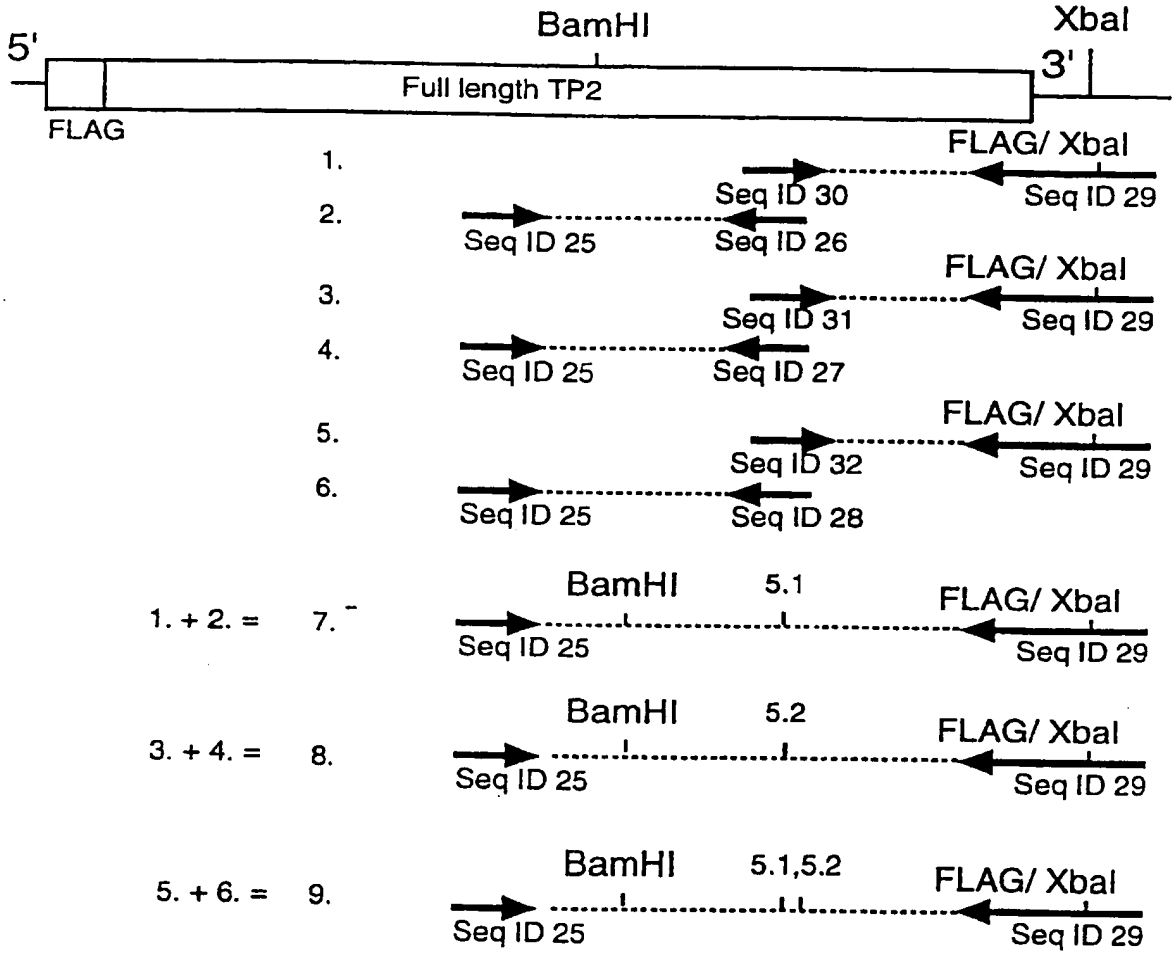
TFNRGFKAGRNMRRKLFGVLRLKCHSLFLDLQVNSLQTVCTNIYKILLL

QAYRFHACVLQLPFHQQVWKNPTFFLRVISDTASLCYSILKAKNAGMSL

GAKGAAGPLPSEAVQWLCHQAFLLKLTRHRVTYVPLLGSLRTAQTQLSR

KLPGTTLTALEAAANPALPSDFKTILD

FIG. 10


5'
BamHI
XbaI
3'
Full length TP2
FLAG
1.
FLAG/ XbaI
Seq ID 30
Seq ID 29
2.
Seq ID 25
Seq ID 26
3.
FLAG/ XbaI
Seq ID 31
Seq ID 29
4.
Seq ID 25
Seq ID 27
5.
FLAG/ XbaI
Seq ID 32
Seq ID 29
6.
Seq ID 25
Seq ID 28
1. + 2. =
7.
BamHI
5.1
FLAG/ XbaI
Seq ID 25
Seq ID 29
3. + 4. =
8.
BamHI
5.2
FLAG/ XbaI
Seq ID 25
Seq ID 29
5. + 6. =
9.
BamHI
5.1,5.2
FLAG/ XbaI
Seq ID 25
Seq ID 29

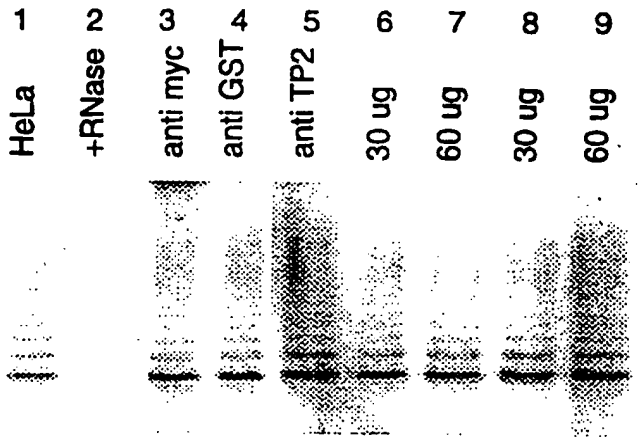


FIG.11A

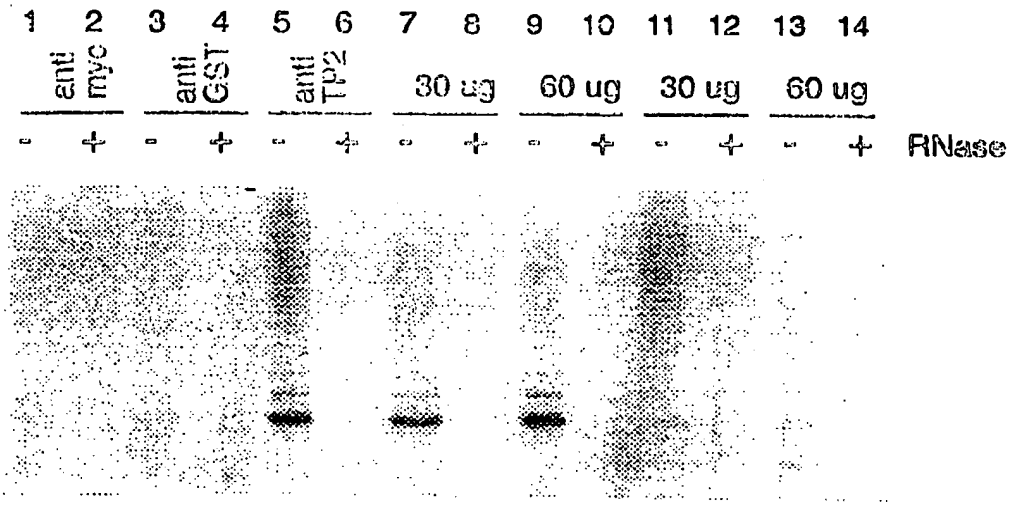


FIG.11B

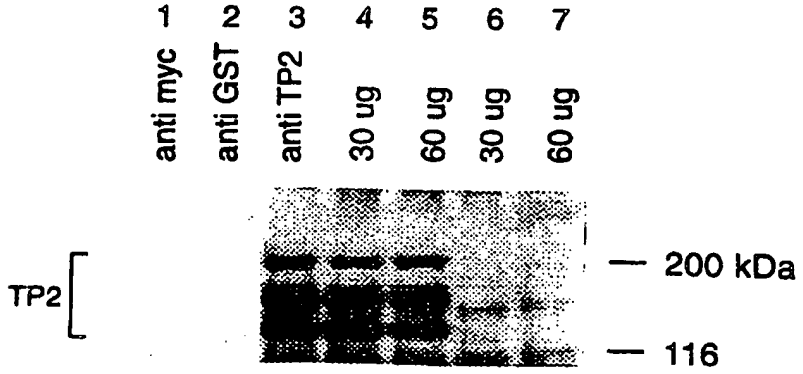


FIG.11C

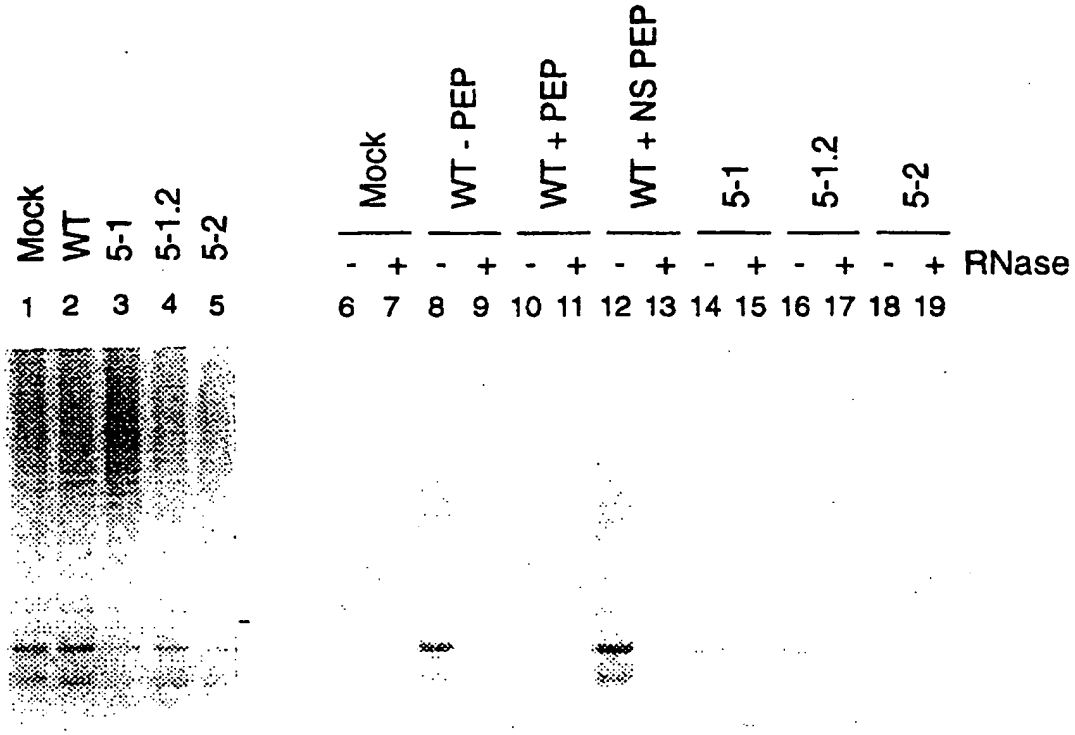


FIG. 12A

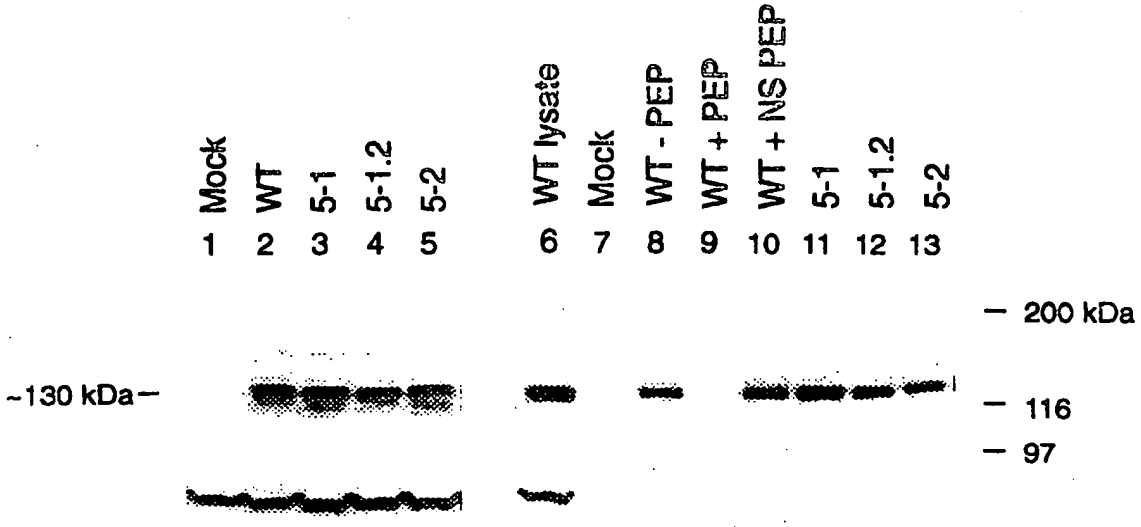


FIG. 12B

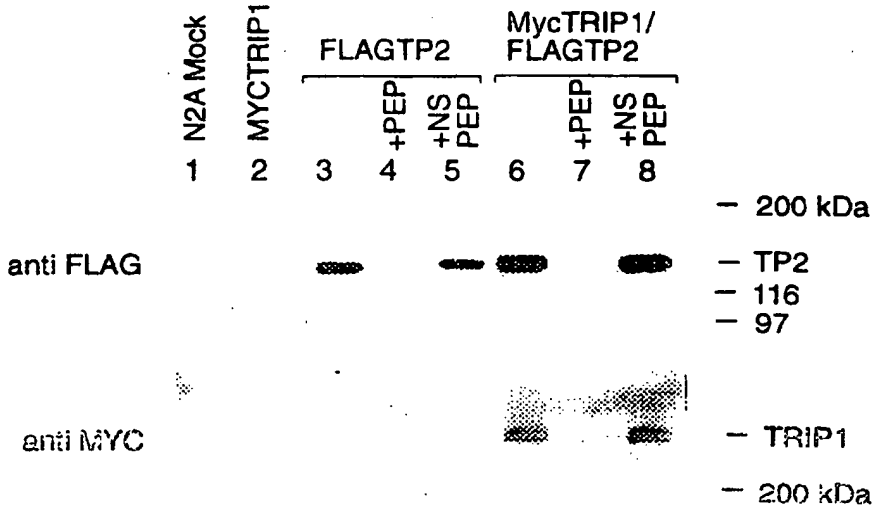


FIG. 13A

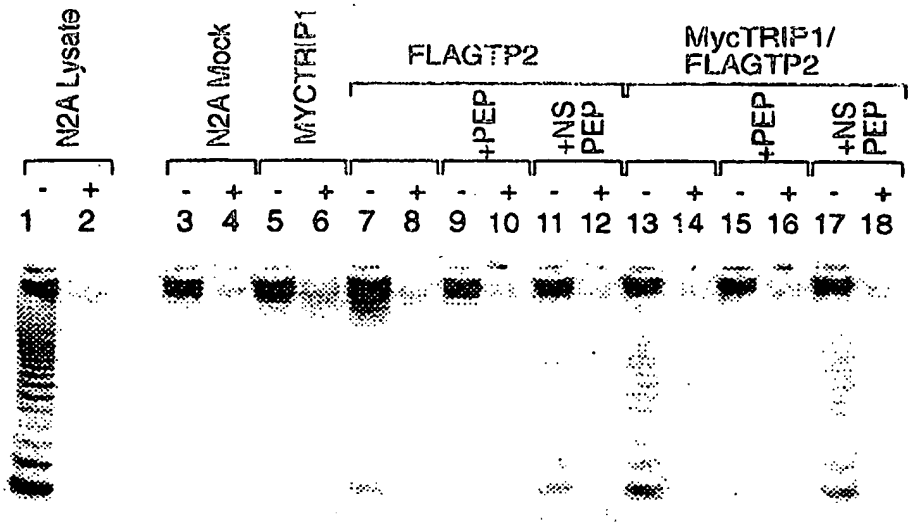


FIG. 13B

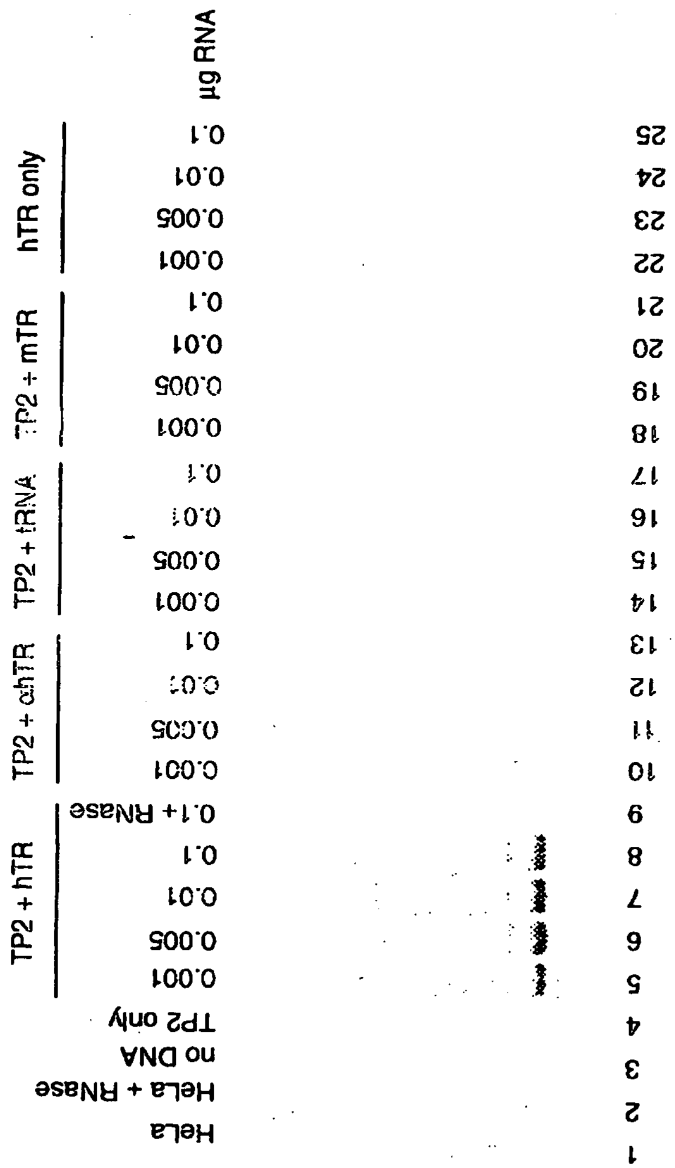

µg RNA
Hela
Hela + RNase
no DNA
TP2 only
TP2 + hTR
0.001
0.005
0.01
0.1
0.1+ RNase
TP2 + chTR
0.001
0.005
0.01
0.1
TP2 + tRNA
0.001
0.005
0.01
0.1
TP2 + mTR
0.001
0.005
0.01
0.1
hTR only
0.001
0.005
0.01
0.1
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25

FIG. 14

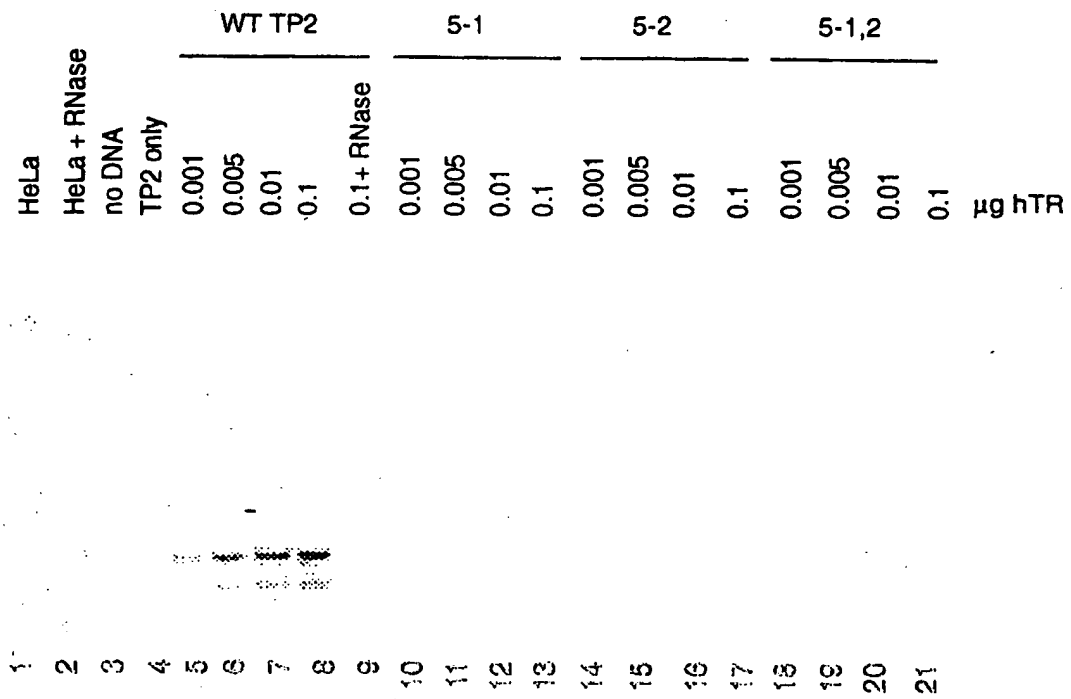


FIG. 15A

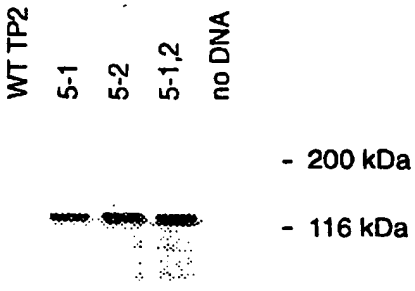


FIG. 15B

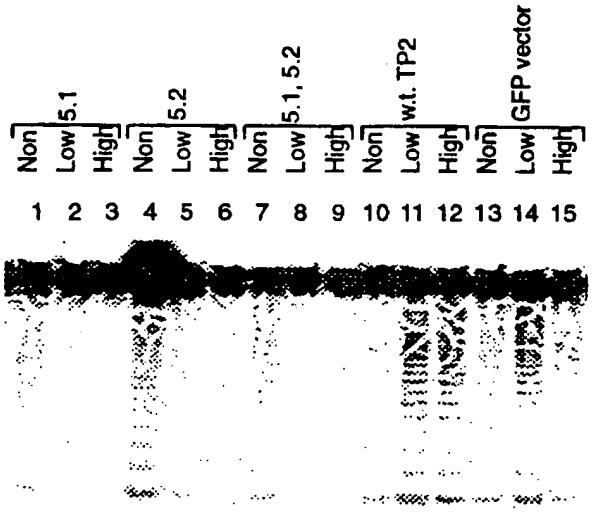


FIG. 16A

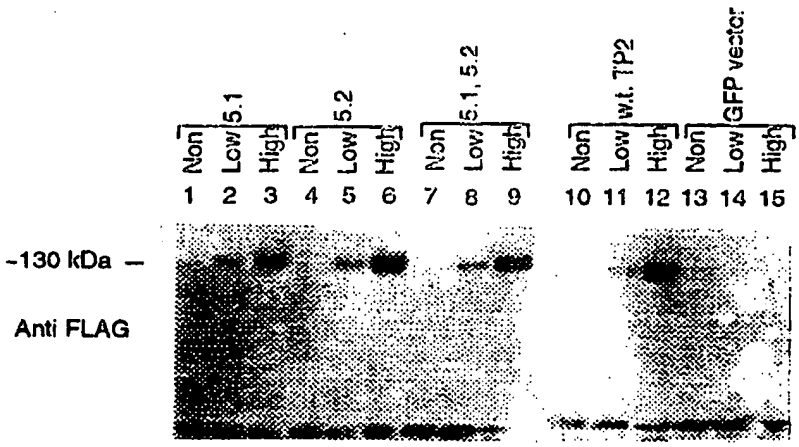


FIG. 16B

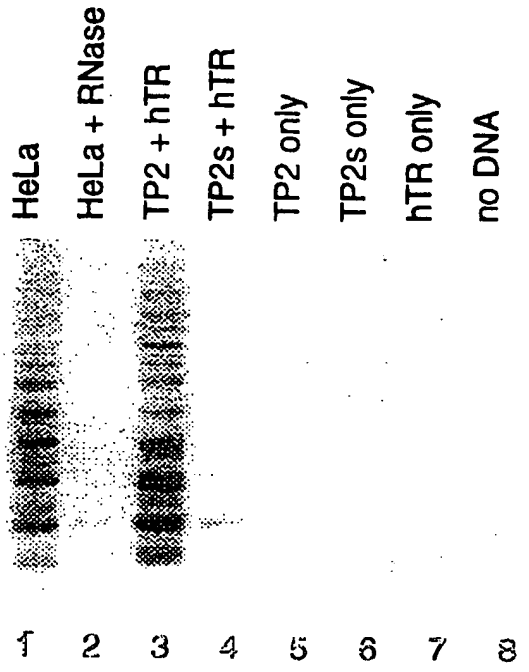


FIG.17A

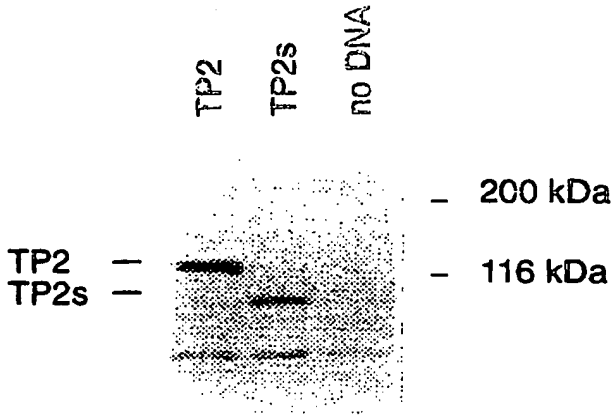


FIG.17B

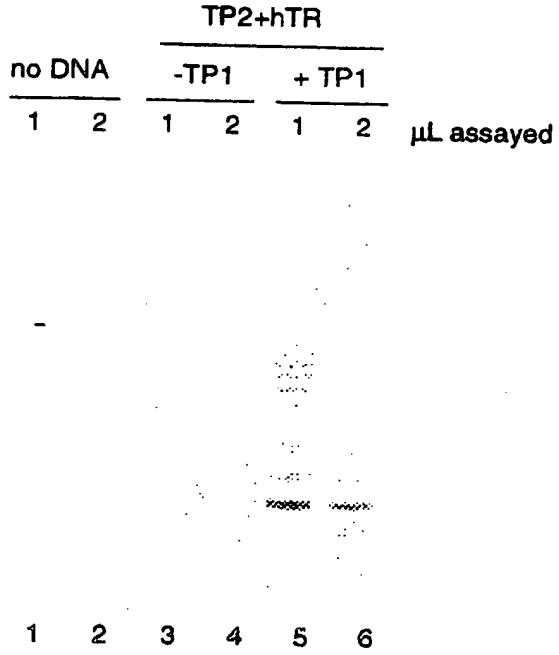


FIG.18